

Zufällige Felder und Perkolationen auf Bäumen

Diplomarbeit

vorgelegt am Fachbereich Mathematik
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

Axel Müller

11. September 1999

Für meine Eltern, ohne deren Liebe und Unterstützung ich niemals so weit gekommen wäre.

Inhaltsverzeichnis

1	Einleitung	5
1.1	Überblick über Inhalt und Ergebnisse	5
1.2	Aufbau der Arbeit	7
2	Zufällige Prozesse auf Bäumen	9
2.1	Bezeichnungen	9
2.2	Elektrische Netzwerke	11
2.3	Perkolationen	16
2.4	Markov-Felder und Gibbs-Maße	17
3	Dominanz und gegenseitige Information	25
4	Das Broadcasting-Modell	35
4.1	Grundlegende Eigenschaften des Modells	35
4.2	Der Interaktionsparameter α	43
4.3	Kritischer Parameter für Informationsfluß	44
4.4	Das Ising-Modell	46
4.5	Kritischer Parameter für Phasenübergang	53
4.6	Vergleich der kritischen Interaktionsstärken	56
4.7	Spinperkolation	59
4.8	Zusammenfassung der Ergebnisse	66
5	Weitreichende Einflüsse	69
5.1	Eine kritische Interaktionsstärke für Informationsfluß	69
5.1.1	Information und effektive Leitfähigkeit	69
5.1.2	Dominanz und unabhängige Version	75
5.1.3	Kritische Interaktion für Informationsfluß	93

5.2	Eine kritische Interaktionsstärke für Phasenübergang	95
6	Anhang	107
6.1	Lyons Nachweis der kritischen Interaktion für Spinperkolation in der Plus-Phase	107
6.1.1	Quasi-Bernoulli-Perkolationen	107
6.1.2	Kritische Interaktion für Spinperkolation	111

Kapitel 1

Einleitung

1.1 Überblick über Inhalt und Ergebnisse

Wir werden uns in dieser Arbeit vorwiegend mit einem Modell befassen, das Y. Peres, C. Kenyon, W. Evans und L.J. Schulman 1998 in ihrem Artikel *“Broadcasting on trees and the Ising-Modell”* eingeführt haben.

In diesem Modell wird ein Signal, das die Werte $+1$ oder -1 annehmen kann, von der Wurzel eines Baumes aus entlang der Äste eines unendlichgroßen Baumes übertragen. Die Kanten des Baumes agieren dabei als Übertragungskanäle zwischen den Knoten. Jede Kante kann das Signal korrekt übertragen oder es flippen, das heißt, das Vorzeichen des Signals umkehren.

Das Übertragungsverhalten der Kanten ist zufällig. Mit einer festen Wahrscheinlichkeit ϵ , mit $0 < \epsilon \leq \frac{1}{2}$, verfälscht eine Kante das Signal. Dies geschieht an allen Kanten unabhängig mit der gleichen Wahrscheinlichkeit. Es stellt sich nun die Frage, wie groß diese Fehlerwahrscheinlichkeit höchstens sein darf, damit das, was in der Krone des Baumes ankommt, noch etwas zu tun hat mit dem, was in der Wurzel eingespeist wird. Mit anderen Worten: Sind die Signale auf Knoten, die einen Abstand $\geq n$ von der Wurzel haben, für $n \rightarrow \infty$ asymptotisch unabhängig vom Signal in der Wurzel? Eine Möglichkeit, den Grad der Abhängigkeit zu messen, ist die sogenannte *Information*, der Kullback-Leibler-Abstand von gemeinsamer Verteilung zur Produkt-Verteilung, die in Definition 16 eingeführt wird.

Wir werden sehen, daß es eine kritische Schwelle $\epsilon_{c,I}$ für Informationsübertragung gibt. Ist die Fehlerwahrscheinlichkeit größer als $\epsilon_{c,I}$, so ist die Informa-

tion, die zwischen Wurzel und Krone übertragen wird, 0. Ist die Fehlerwahrscheinlichkeit kleiner als $\epsilon_{c,I}$, so wird Information übertragen. Dieser kritische Wert $\epsilon_{c,I}$ hängt nur von der Branching-Number, einer Art *mittleren* Verzweigungszahl, des Baumes (vgl. Definition 1) ab.

Wir werden sehen, daß das Broadcasting-Modell eine elegante Formulierung eines wohlbekannten Modells, des Ising-Modells, mit freien Randbedingungen, ist.

Im Ising-Modell hat jeder Knoten des Baumes einen “magnetischen” Spin, der entweder $+1$ oder -1 sein kann. Spins direkt benachbarter Knoten beeinflussen sich, in dem sie versuchen, den gleichen Wert anzunehmen. Diesem Effekt wirkt ein thermischer Einfluß entgegen, der mittels eines als *Temperatur* bezeichneten Parameters modelliert wird.

Die klassische Frage im Ising-Modell ist, ob Phasenübergang stattfindet. Wir wollen Phasenübergang als das Phänomen verstehen, daß die Wurzel des Baumes die Vorgabe von Randbedingungen auf der Krone des Baumes spürt. Ist dies der Fall, so sagen wir, daß *Phasenübergang stattfindet*. Auch dies ist eine Form der gegenseitigen Beeinflussung zwischen Wurzel und Krone des Baumes. Russel Lyons hat 1989 in seinem Artikel “*The Ising-Model on trees and tree-like Graphs*” das Ising-Modell auf Bäumen untersucht und gezeigt, daß es eine kritische Temperatur t_c für Phasenübergang gibt. Ist die Temperatur höher als t_c , so spürt die Wurzel nichts von den Randbedingungen der Krone; ist die Temperatur geringer als t_c , so haben die Randbedingungen Einfluß auf die Wurzel. Auch hier hängt die kritische Temperatur nur von der Branching-Number des Baumes ab.

In der Broadcasting-Formulierung des Modells ist der Fluß von Information ein naheliegendes Werkzeug, um die Beeinflussung von Wurzel und Krone zu messen, in der Ising-Formulierung ist die Existenz von Phasenübergang ein ebenso naheliegendes Werkzeug, ebendiesen Einfluß zu messen.

Wir werden die beiden Arten der Beeinflussung miteinander vergleichen und können zeigen, daß für die Übertragung von Information stets eine stärkere Interaktion zwischen den Knoten notwendig ist, als für den Einfluß der Randbedingungen aus der Krone.

Als letztes Phänomen werden wir untersuchen, ob es einen Pfad im Baum gibt, der in der Wurzel startend nur Knoten gleichen Spins besucht und die unendlich weit entfernte Krone erreicht. Wir bezeichnen dieses Phänomen als

Spinperkolation.

Wir werden die Berechnung der kritischen Interaktion für Spinperkolation in einem Bernoulli-Feld auf den Kanten rekapitulieren und dann zeigen, daß die Existenz eines Perkulationspfades nur von der Interaktionsstärke des Modells und nicht von etwaigen Randbedingungen abhängt. Dabei kombinieren wir Ergebnisse aus zwei Arbeiten von Lyons und die Erkenntnis, daß Broadcasting-Modell und freies Ising-Modell identisch sind. Wir erhalten so einen neuen, einfachen Beweis über die kritische Interaktion für Spinperkolation in der Plus-Phase des Ising-Modells, die Lyons bereits in [7] berechnet hat.

1.2 Aufbau der Arbeit

In Kapitel 1 beschreiben wir kurz die Strukturen und Begriffe, die wir in den späteren Beweisen verwenden werden. Es handelt sich dabei um Anleihen aus den elektrischen Netzwerken auf Bäumen, Perkulationen auf Bäumen und den Gibbs-Maßen und Markov-Feldern. Wir zitieren einige Ergebnisse, auf die wir uns später beziehen werden.

Das zweite Kapitel stellt einige Hilfsmittel aus der Informationstheorie zusammen. Wir führen in Definition 16 den zentralen Begriff der gegenseitigen Information von Zufallsvariablen ein, der als ein Abstandsbegriff für zwei Verteilungen verstanden werden kann. Wir werden die gegenseitige Information mit einem zweiten Abstandsbegriff, einer symmetrisierten Form des χ^2 Abstands, vergleichen.

In Definition 17 werden wir einen Dominanzbegriff für Paare von Zufallsvariablen einführen, der sich an dem von Peres et al in [11] für Bäume eingeführten Dominanzbegriff orientiert und der es uns ermöglichen wird, die gegenseitige Information zwischen Zufallsvariablen abzuschätzen.

Der Kern der Arbeit befindet sich in Kapitel 3. Hier wird das Broadcasting-Modell eingeführt und charakterisiert. Wir zeigen in Satz 12, daß es eine kritische Interaktion für Information zwischen Spin der Wurzel und den Spins der Krone gibt. Weiterhin zeigen wir, daß Broadcasting-Modell und Ising-Modell mit freien Randbedingungen gleich sind, wenn man die Modellparameter geeignet ineinander überführt. Wir berechnen in Satz 15 die kritische Interaktion für Phasenübergang und vergleichen diese in Satz 16 mit der kritischen Interaktion für Informationsfluß.

Schließlich betrachten wir das Phänomen der Spinperkolation und zeigen, daß es auch hier eine kritische Interaktion gibt. Die Aussagen aus Korollar 8 und Satz 19 zeigen, daß Spinperkolation nicht von etwaigen Randbedingungen abhängt. Der Beweis von Satz 19 benutzt einige Eigenschaften des Broadcasting-Modells und ist einfacher als der in [7], wo die gleiche Aussage für die Ising-Formulierung des Modells gezeigt wird.

Im vierten Kapitel beweisen wir die Sätze, die Aussagen über die kritischen Parameter für Phasenübergang und Informationsfluß machen. Dazu geben wir in Satz 20 eine untere Abschätzung der Information zwischen Spin der Wurzel und den Spins der Krone mittels des effektiven Widerstandes eines Netzwerkes an.

Satz 24 bietet eine obere Abschätzung der Information zwischen Wurzel und Krone, durch die Information im Broadcasting-Modell auf einer *unabhängigen Version* des Baumes.

Der Beweis von Satz 15, der eine Aussage über die kritische Interaktion für Phasenübergang macht, basiert auf der Beobachtung, daß wir den Einfluß der Plus-Randbedingungen auf einen Knoten durch den Einfluß auf seine Nachfolger ausdrücken können. Dies ermöglicht es uns dann, Flüsse zu konstruieren, die den Einfluß des Randes auf die Wurzel abschätzen.

Im Anhang stellen wir noch den ursprünglichen Beweis von Lyons über eine kritische Interaktion für Spinperkolation in der Plus-Phase dar, der keine Aussagen über das freie Ising-Modell verwendet.

Kapitel 2

Zufällige Prozesse auf Bäumen

2.1 Bezeichnungen

Mit $T = (K_T, E_T)$ wird, sofern nicht ausdrücklich anders gesagt, stets ein verwurzelter, lokal endlicher Baum mit Knotenmenge K_T und Kantenmenge E_T bezeichnet. Lokal endlich bedeutet hier, daß jeder Knoten $\sigma \in T$ nur endlich viele Nachbarn hat. Es gebe einen ausgezeichneten Knoten in T , die Wurzel des Baumes, die mit r bezeichnet wird. σ, τ, ξ sind stets Bezeichnungen für Knoten des Baumes, Kanten werden mit e_1, e_2 , usw. benannt. Für zwei benachbarte Knoten σ und τ sei $e_{(\sigma, \tau)}$ die Kante zwischen σ und τ .

Für $A \subset K_T$ definieren wir den **Rand** ∂A von A als:

$$\partial A := \{\tau \in (K_T \setminus A) : \exists \sigma \in A \text{ mit } e_{(\sigma, \tau)} \in E_T\}$$

Die Menge der endlichen Teilmengen von K_T bezeichnen wir mit $\mathcal{C}(T)$. Für eine endliche Teilmenge B sei $\#B$ die Mächtigkeit der Menge.

Eine in r startende maximale Folge benachbarter Knoten, in der kein Knoten mehrfach vorkommt, bezeichnen wir als **Ast**. Sofern nicht anders gesagt, werden wir uns stets mit Bäumen befassen, die keine endlichen Äste haben.

Ist T endlich, so bezeichnen wir die **Krone** von T als ∂T . Die Krone ist dabei die Menge aller Knoten in T , die, von der Wurzel aus gesehen, keine Nachfolger in T haben. Einen Knoten der Krone bezeichnen wir als **Blatt**.

Für $\sigma \in K_T, \sigma \neq r$ ist $\overleftarrow{\sigma}$ der eindeutig bestimmte **Vorgänger** des Knotens σ , also der Knoten, der auf dem Pfad von der Wurzel zu σ unmittelbar vor σ besucht wird. $\overrightarrow{\sigma}$ bezeichnet in ebendiesem Sinne die Menge aller **Nachfolger** des Knotens σ . Die direkten **Nachbarn** eines Knotens σ bezeichnen wir als $N(\sigma) := \{\overleftarrow{\sigma}\} \cup \overrightarrow{\sigma}$.

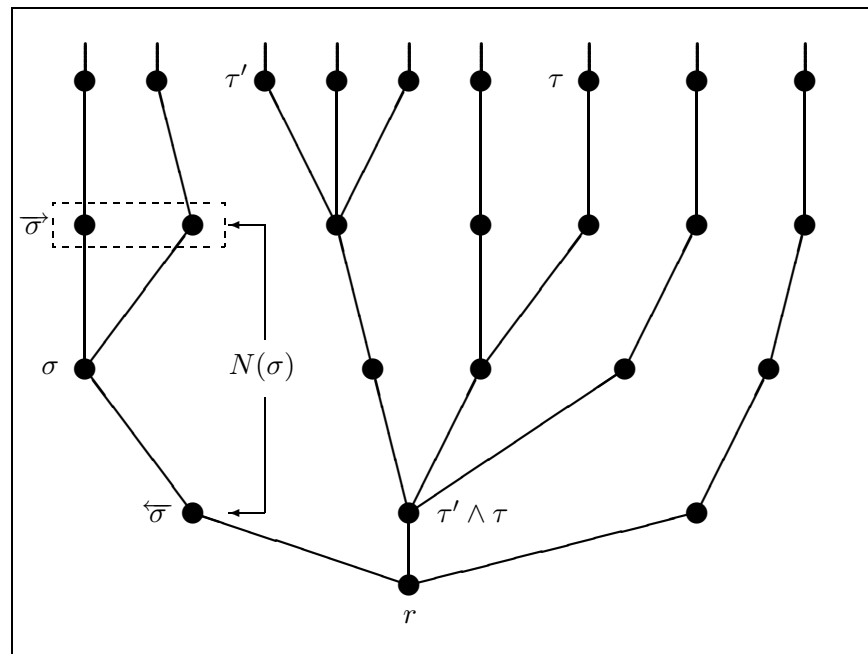


Abbildung 2.1: Ein Baum T mit einigen der benutzten Bezeichnungen

Auf T wird im kanonischen Sinne der Abstand eines Knotens σ von der Wurzel r definiert. Für $\sigma \in T, \sigma \neq r$ ist $|\sigma| := |\text{Pfad}(\sigma)| - 1$. Den Abstand einer Menge $A \subset K_T$ von der Wurzel definieren wir auf naheliegende Weise als:

$$|A| := \inf \sigma \in A|\sigma|$$

Für Kanten $e \in E_T$ definieren wir den Abstand von der Wurzel als:

$$|e_{(\overleftarrow{\sigma}, \sigma)}| = |\overleftarrow{\sigma}|$$

Mit diesem Abstandsbegriff läßt sich die Menge der Nachbarn von σ als

$$N(\sigma) := \{\tau \in K_T : ||\sigma| - |\tau|| = 1\} \quad (2.1)$$

schreiben. $\hat{N}(\sigma)$ sei $N(\sigma) \cup \{\sigma\}$.

Eine Menge $\Pi \subset K_T$ für die gilt, daß jeder Ast genau einen Knoten mit Π gemeinsam hat, wird als **Cutset** bezeichnet.

Eine spezielle Art von Cutsets sind die Sphären. Für $n \in \mathbb{N}$ wird die **n-Sphäre** von T definiert als:

$$T^n := \{\sigma \in T \mid |\sigma| = n\}$$

Zu jedem Cutset Π ist das Innere $(< \Pi) \subset K_T$ definiert als die Menge der Knoten, die zwischen der Wurzel r und den Knoten von Π liegen. $(\leq \Pi)$ wird auf naheliegende Weise als $(\leq \Pi) := (< \Pi) \cup \Pi$ definiert. Für $\sigma \in K_T$ ist $T_\sigma \subset T$ der Teilbaum von T mit Wurzel σ .

Die wichtigste Kenngröße eines Baumes T im Rahmen des stochastischen Modells, das hier betrachtet wird und weit darüber hinaus ist die *Branching-Number* $\text{br}T$. Sie wurde von Russel Lyons in [7] eingeführt und kann als mittlere Kantenzahl pro Knoten verstanden werden. Wir geben hier eine Definition unter Verwendung von Cutset-Summen an und beschreiben später einige wichtige Zusammenhänge zu anderen Prozessen auf Bäumen.

Definition 1. *Gegeben sei ein lokal endlicher Baum T . Die **Branching-Number** $\text{br}T$ ist definiert als*

$$\text{br}T := \sup \left\{ \lambda \geq 1 : \inf_{\Pi \subset K_T} \sum_{\tau \in \Pi} \lambda^{-|\tau|} > 0 \right\} \quad (2.2)$$

2.2 Elektrische Netzwerke

Ein Modell, aus dem wir einige Anleihen machen werden, ist das der elektrischen Netzwerke auf Graphen, wie sie zum Beispiel in [2] und [9] betrachtet werden. Wir führen die später auftretenden Größen ein und zitieren die Aussagen, die wir verwenden werden. Da wir uns nicht mit Irrfahrten auf Graphen

beschäftigen werden, verzichten wir weitgehend auf eine wahrscheinlichkeitstheoretische Interpretation der Größen. Hierfür sei auf die oben erwähnten Arbeiten, insbesondere auf das vorzügliche Buch von Doyle und Snell, verwiesen. Wir beschränken uns weitgehend auf den Spezialfall, daß der zugrundeliegende Graph ein Baum ist.

Unter einem elektrischen Netzwerk wollen wir einen Baum T mit einer positiven Gewichtung seiner Kanten verstehen. Die Gewichte der Kanten verstehen wir als elektrische Leitfähigkeiten der Kanten. Wir definieren also:

Definition 2. *Gegeben sei ein lokal endlicher Baum T . Weiterhin sei jeder Kante $e \in E_T$ eine als **Leitfähigkeit** bezeichnete positive, reelle Zahl $C(e)$ zugeordnet. $R(e) := 1/C(e)$ heißt **Widerstand** der Kante e . Die Menge aller Leitfähigkeiten sei $C := \{C(e) : e \in E_T\}$. Dann heißt (T, C) **elektrisches Netzwerk** auf T .*

Wir führen nun die Begriffe Strom i und Spannung v ein und lassen uns dabei ganz von der physikalischen Anschauung leiten.

Seien A und Z zwei disjunkte Mengen von Knoten aus T . Wir stellen uns nun vor, daß alle Knoten in A mit dem Plus-Pol einer Batterie, alle Knoten in Z mit dem Minus-Pol einer Batterie verbunden werden. Dies führt dazu, daß sich in den Knoten zwischen A und Z eine Spannung v einstellt und in den Kanten zwischen A und Z ein Strom i fließt. Strom und Spannung erfüllen dabei folgende Gesetze.

Ohmsches Gesetz

Falls die Knoten τ und σ benachbart sind, gilt:

$$v(\sigma) - v(\tau) = R(e_{(\sigma,\tau)}) \cdot i(\sigma, \tau)$$

Kirchhoffsches Gesetz

Für alle Knoten $\sigma \notin A \cup Z$ gilt:

$$i(\overleftarrow{\sigma}, \sigma) + \sum_{\tau \in \overrightarrow{\sigma}} i(\sigma, \tau) = 0$$

Diese beiden Gesetze, zusammen mit den durch den Batterie-Anschluß gegebenen Randbedingungen, legen die Funktionen Spannung $v : K_T \rightarrow \mathbb{R}_{+,0}$ und Strom $i : K_T \times K_T \rightarrow \mathbb{R}$ eindeutig fest. Beide haben, wenn man Irrfahrten auf

T betrachtet, eine wichtige Interpretation. Wir verweisen wieder die ausführliche Darstellung in [9] und [2].

Die hier auftauchende Abbildung i ist ein Fluß auf T . Flüsse werden in den später auftauchenden Beweisen eine wichtige Rolle spielen. Wir definieren daher ganz allgemein:

Definition 3. Seien $A, Z \subset K_T$ zwei disjunkte Mengen. Wir nennen eine Abbildung $j : K_T \times K_T \rightarrow \mathbb{R}$ einen **Fluß** von A nach Z , wenn für alle $\sigma, \tau \in K_T$ gilt:

- $j(\sigma, \tau) = -j(\tau, \sigma)$
- $j(\overleftarrow{\sigma}, \sigma) + \sum_{\tau \in \overrightarrow{\sigma}} j(\sigma, \tau) = 0$, falls $\sigma \neq r$
- $j(\sigma, \tau) = 0$, falls $\tau \notin N(\sigma)$

Wir setzen j auf K_T fort:

$$j(\sigma) := \sum_{\tau \in \overrightarrow{\sigma}} j(\sigma, \tau)$$

Die Größe

$$\text{Strength}(j) := \sum_{\sigma \in A} j(\sigma) \quad (2.3)$$

nennen wir die **Stärke** des Flusses j . Ein Fluß der Stärke 1 heißt **unit flow**.

Im weiteren Verlauf betrachten wir meist Flüsse, die von der Wurzel des Baumes zur Krone gehen. Da hierdurch eine Richtung ausgezeichnet ist, schreiben wir für $j(\overleftarrow{\sigma}, \sigma)$ häufig $j(e_{(\overleftarrow{\sigma}, \sigma)})$ oder $j(e)$. Mit dieser Vereinbarung können wir kürzer formulieren, welche charakteristische Eigenschaft ein Fluß hat:

$j : T \rightarrow \mathbb{R}$ ist ein **Fluß** auf T , wenn für alle $\sigma \in K_T$ gilt:

$$j(\sigma) = \sum_{\tau \in \overrightarrow{\sigma}} j(\tau) \quad (2.4)$$

Ist j ein Fluß von der Wurzel r in die n -Sphäre T^n , so gilt für alle Cutsets Π , die zwischen r und T^n liegen:

$$\sum_{\sigma \in \Pi} j(\sigma) = j(r) = \text{Strength}(j) \quad (2.5)$$

Wie bereits erwähnt, erzeugt eine Spannungsverteilung v mit $v|_A = v_a$ (dabei sei hier und im folgenden $f|_A$ die Einschränkung von f auf A) und $v|_Z = 0$ einen Strom i von A nach Z . i ist ein Fluß und erfüllt i das Ohmsche und das Kirchhoffsche Gesetz. Wenn wir v_a nun so wählen, daß i die Stärke 1 hat, so heißt i der **unit current flow**.

Eine wichtige Größe in einem elektrischen Netzwerk ist der effektive Widerstand. Wir definieren diesen Begriff, dem physikalischen Pendant folgend, als:

Definition 4. Sei $T; C$ ein elektrisches Netzwerk und seien $A, Z \subset K_T$ zwei disjunkte Mengen. v sei eine Spannungsverteilung mit $v|_A = v_a$ und $v|_Z = 0$. Weiterhin mögen v und der von ihr erzeugte Strom i dem Ohmschen und dem Kirchhoffschen Gesetz genügen. Dann ist

$$C_{\text{eff}}(A, Z) := \frac{\sum_{\sigma \in A} i(\sigma)}{v_a} \quad (2.6)$$

die **effektive Leitfähigkeit** zwischen A und Z . Ihr Kehrwert R_{eff} heißt **effektiver Widerstand** zwischen A und Z .

Wir werden uns hauptsächlich mit Netzwerken befassen, bei denen ein Strom von der Wurzel r zur unendlich weit entfernten Krone fließt. Wir definieren die effektive Leitfähigkeit als einen Grenzwert:

Definition 5. Sei T ein lokal endlicher Baum ohne endliche Äste. (T, C) sei ein elektrisches Netzwerk und $\Pi_n \subset K_T$ sei eine Folge von Cutsets mit $|\Pi_n| \rightarrow \infty$. v_n sei eine Folge von Spannungsverteilungen, die den Randbedingungen $v_n(r) = v_r, v_n|_{\Pi_n} = 0$ genügen. Dann heißt

$$C_{\text{eff}}((T, C)) := C_{\text{eff}}(r, \partial T) := \lim_{n \rightarrow \infty} C_{\text{eff}}(r, \Pi_n) \quad (2.7)$$

effektive Leitfähigkeit von (T, C) . Ihr Kehrwert R_{eff} ist der effektive Widerstand von (T, C) .

Es besteht ein fundamentaler Zusammenhang zwischen der effektiven Leitfähigkeit eines sehr einfach konstruierten Netzwerkes auf T und der Branching-Number von T . Wir werden diesen Zusammenhang später wesentlich verwenden:

Satz 1 ([9]). *Sei T ein lokal endlicher Baum ohne endliche Äste und sei $\lambda \geq 1$ eine reelle Zahl. Jeder Kante e von T werde die Leitfähigkeit $\lambda^{-|e|}$ zugewiesen. Das so entstandene Netzwerk wird kurz mit T_λ bezeichnet. Dann gilt:*

$$\sup\{\lambda \geq 1 : R_{\text{eff}}(T_\lambda) < \infty\} = \text{br}T \quad (2.8)$$

Ein durchaus einleuchtendes Ergebnis im Zusammenhang mit der effektiven Leitfähigkeit eines Netzwerkes ist das Rayleigh-Prinzip. Es besagt, daß die Erhöhung der Leitfähigkeiten von einzelnen Kanten die effektive Leitfähigkeit des Netzwerkes höchstens erhöhen kann.

Satz 2 ([2] Rayleigh-Prinzip). *Seien (T, C) und (T, \tilde{C}) zwei elektrische Netzwerke auf T mit der Eigenschaft $C(e) \geq \tilde{C}(e)$ für alle $e \in E_T$. Dann gilt:*

$$\begin{aligned} C_{\text{eff}}((T, C)) &\geq C_{\text{eff}}((T, \tilde{C})) \\ \iff R_{\text{eff}}((T, C)) &\leq R_{\text{eff}}((T, \tilde{C})) \end{aligned}$$

Als letzten Begriff wollen wir die Energie eines Flusses einführen. Wir definieren diese so wie die in einem physikalischen Netzwerk umgesetzte Wärmedissipation.

Definition 6. *Gegeben sei ein elektrisches Netzwerk (T, C) und ein Fluß j von A nach Z . Dann ist die **Energie** von j definiert als:*

$$E(j) := \sum_{e \in E_T} j(e)^2 \cdot R(e) \quad (2.9)$$

Wir haben den Fluß mit Stärke 1, der dem Ohmschen und Kirchhoffschen Gesetz genügt, als unit current flow eingeführt. Der folgende Satz zeigt, daß dieser Fluß unter allen unit flows die geringste Energie hat.

Satz 3 (Thompson-Prinzip, [2]). *Sei (T, C) ein elektrisches Netzwerk und sei i der unit current flow von A nach Z . j sei ein anderer unit flow von A nach Z . Dann gilt:*

$$E(i) < E(j) \quad (2.10)$$

Eine wichtige Frage ist, ob die Energie eines Flusses endlich ist. Als letztes Ergebnis geben wir einen Satz an, der hierzu eine Aussage macht.

Satz 4 ([9], Proposition 14). *Sei (T, C) ein elektrisches Netzwerk und seien w_n positive reelle Zahlen mit $\sum_{n \geq 1} w_n < \infty$. Weiterhin sei j ein Fluß auf T mit der Eigenschaft $j(e) \leq w_{|e|}$. Dann hat j endliche Energie:*

$$E(j) = \sum_{e \in E_T} j(e)^2 R(e) < \infty$$

2.3 Perkolationen

Ein elementarer Prozeß auf einem Baum ist der der **Perkolation**. Wir werden uns hier ausschließlich mit Bond-Perkolationen befassen. Man versteht anschaulich unter einer Bond-Perkolation eines Baumes T das Entfernen von Kanten.

Bei unendlichen Bäumen stellt sich im Zusammenhang mit einer Perkolation die natürliche Frage, ob es im perkolierten Baum noch unendlich große Zusammenhangskomponenten gibt, insbesondere, ob es noch einen unendlichen Pfad von der Wurzel aus gibt. Existiert ein solcher Pfad, so sagt man, daß *Perkolation stattfindet*. Wir definieren also:

Definition 7. *Sei T ein Baum. $\mathcal{T} := \{(K_T, \hat{E}) : \hat{E} \subset E_T\}$ sei die Menge aller Teilgraphen von T , die alle Knoten aus T haben. \mathbf{P} sei ein Wahrscheinlichkeitsmaß auf \mathcal{T} . Dann heißt $(\mathcal{T}, \mathbf{P})$ eine **Perkolation** von T .*

Für $\sigma \in K_T$ sei \hat{T}_σ die Zusammenhangskomponente von σ in \hat{T} . Weiterhin sei für $\sigma, \tau \in K_T$ das Ereignis $(\sigma \leftrightarrow \tau)$ definiert als:

$$(\sigma \leftrightarrow \tau) := \{\hat{T} : \sigma \in \hat{T}, \tau \in \hat{T}\}$$

Wie bereits gesagt, können wir uns ein $\hat{T} \in \mathcal{T}$ immer so vorstellen, daß in T (zufällig) Kanten entfernt werden.

Eine der einfachsten Perkolationen ist die *Bernoulli-Perkolation*, bei der jede Kante unabhängig mit gleicher Wahrscheinlichkeit entfernt wird.

Definition 8. *Wir nennen eine Perkolation $(\mathcal{T}, \mathbf{P}_{p_s})$, bei der jede Kante mit Wahrscheinlichkeit $1 - p_s$ unabhängig von allen anderen Kanten entfernt wird, eine **Bernoulli-Perkolation** auf T mit **Überlebenswahrscheinlichkeit** p_s .*

Ein elementarer Zusammenhang zwischen einer Bernoulli-Perkolation auf einem Baum und dessen Branching-Number wurde in [7] hergestellt. Es ist der folgende

Satz 5 ([7], Theorem 6.2). *Sei T ein lokal endlicher Baum ohne endliche Äste und seien $(T, \mathbf{P}_{p_s}), p_s \in \mathbb{R}$, eine Familie von Bernoulli-Perkolationen auf T mit den Überlebenswahrscheinlichkeiten p_s . Weiterhin sei:*

$$p_c := p_c(T) = \sup\{p_s : \mathbf{P}_{p_s}[\hat{T}_r = \infty] = 0\}$$

Dann gilt:

$$p_c(T) = \frac{1}{\text{br}T} \quad (2.11)$$

2.4 Markov-Felder und Gibbs-Maße

Wir fassen hier einige Begriffe und Aussagen über Gibbs-Maße und Markov-Felder zusammen, die aus [10] und [6] entnommen wurden.

Im Rahmen dieser Arbeit werden wir uns im wesentlichen mit zufälligen Belegungen auf den Knoten eines Baumes T beschäftigen. Dabei wird jedem Knoten $\sigma \in T$ eine als **Spin** bezeichnete Zufallsvariable $u_\sigma \in \{-1, +1\}$ zugeordnet. Für $A \subseteq K_T$ bezeichnen wir die Menge aller möglichen Belegungen als $\mathcal{U}^A := \{-1, +1\}^A$.

Für $A \subset B \subset K_T$ sei $[A, B]$ definiert als die Menge der Belegungen auf T , die auf A den Wert $+1$ annehmen und auf $(B \setminus A)$ den Wert -1 annehmen:

$$[A, B] := \{a \in \mathcal{U}^{K_T} : a|_A \equiv +1, a|_{(B \setminus A)} \equiv -1\}$$

Weiterhin sei $[A]$ die Menge der Konfigurationen, die auf A den Wert $+1$ annehmen:

$$[A] := \{a \in \mathcal{U}^{K_T} : a|_A \equiv +1\}$$

Ist $B \in \mathcal{C}(T)$, so bezeichnen wir $[A, B]$ als endlichdimensionalen Zylinder. Für $a \in \mathcal{U}^A, b \in \mathcal{U}^B$ schreiben wir kurz (a, b) für die Konfiguration aus $\mathcal{U}^A \times \mathcal{U}^B$ mit $(a, b)|_A = a$ und $(a, b)|_B = b$

Wir werden nun \mathcal{U}^{K_T} eine topologische Struktur geben. Dazu versehen wir

zunächst $\{-1, +1\}$ mit der diskreten Topologie und $\mathcal{U}^{K_T} = \{-1, +1\}^{K_T}$ mit der Produkttopologie. In der Produkttopologie sind die endlichdimensionalen Zylinder offen, sie bilden sogar eine Basis offener Mengen. Mit \mathcal{A} bezeichnen wir die von den endlichdimensionalen Zylindern erzeugte σ -Algebra auf \mathcal{U}^{K_T} . Wie üblich ist $C(\mathcal{U}^{K_T})$ die Menge der stetigen Abbildungen von \mathcal{U}^{K_T} nach \mathbb{R} . $\mathcal{P}(T)$, die Menge der Wahrscheinlichkeitsmaße auf $(\mathcal{U}^{K_T}, \mathcal{A})$, ist eine Teilmenge des Dualraums von $C(\mathcal{U}^{K_T})$. Wir geben $\mathcal{P}(T)$ eine Topologie, nämlich die Relativtopologie des Dualraums, die vage Topologie. Für Maße $\mathbf{P}^n, \mathbf{P} \in \mathcal{P}(T)$ gilt:

$$\mathbf{P}^n \longrightarrow \mathbf{P} \text{ vage} \iff \forall A \in \mathcal{C}(T), a \in \mathcal{U}^A : \mathbf{P}^n[u_A = a] \longrightarrow \mathbf{P}[u_A = a]$$

Wir werden uns nur mit einer Klasse von Wahrscheinlichkeitsmaßen auf $(\mathcal{U}^{K_T}, \mathcal{A})$ befassen, den Gibbs-Maßen und Markov-Feldern. Um die Definition eines Gibbs-Maßes formulieren zu können, müssen wir zunächst sagen, was wir unter einem **Potential** auf \mathcal{U}^{K_T} verstehen wollen.

Satz und Definition 1. *Wir nennen eine Abbildung V , die für alle $A \in \mathcal{C}(T)$ jedem $a \in \mathcal{U}^A$ eine reelle Zahl $V(a)$ zuordnet, ein **Potential** auf \mathcal{U}^{K_T} .*

*Zu jedem Potential V gibt es ein ihm zugehöriges **Interaktionspotential** I_V auf \mathcal{U}^{K_T} . Für $A \subset \mathcal{C}(T), a \in \mathcal{U}^A$ ist $I_V(a)$ definiert als:*

$$I_V(a) := \sum_{X \subset A} (-1)^{|A \setminus X|} V(a|_X)$$

$V(a)$ kann mit I_V dargestellt werden:

$$V(a) = \sum_{X \subset A} I_V(a|_X)$$

Falls ein Potential V die Eigenschaft erfüllt

$$I_V(a) \neq 0 \implies \exists \sigma \in K_T \text{ mit } a \in \{-1, +1\}^{\{\sigma, \bar{\sigma}\}},$$

*so heißt V (**nächstes**) **Nachbarschaftspotential** auf \mathcal{U}^{K_T} .*

Das Potential V modelliert die Interaktion zwischen den Spins der Knoten. Das ihm zugeordnete Interaktionspotential gibt an, welche Teile eines Bereichs miteinander interagieren. Tragen nur Paare von benachbarten Spins zum Potential einer Konfiguration bei, so heißt V nächstes Nachbarschaftspotential. Wir werden uns ausschließlich mit Potentialen dieser Art beschäftigen.

Nun können wir sagen, was wir unter einem Gibbs-Zustand verstehen wollen.

Definition 9. Sei $A \subset \mathcal{C}(T)$, $b \in \mathcal{U}^{\partial A}$ und V ein Nachbarschaftspotential auf K_T . Wir nennen $F_A^b : \mathcal{U}^A \rightarrow [0, 1]$ mit

$$F_A^b(a) := \frac{\exp[V((a, b))]}{\sum_{c \in \mathcal{U}^A} \exp[V((c, b))]}, \quad a \in \mathcal{U}^A$$

Gibbs-Zustand auf A mit Randbedingung b und Potential V .

Wir werden diesen Begriff verwenden, um zu definieren, was ein Gibbs-Maß auf T ist:

Definition 10. Sei V ein Nachbarschaftspotential auf T . Ein Wahrscheinlichkeitsmaß $\mathbf{P} \in \mathcal{P}(T)$ heißt **Gibbs-Maß mit Potential V** , falls gilt:

- i) Für alle $A \in \mathcal{C}(T)$, $a \in \mathcal{U}^A$ ist $\mathbf{P}[u_A = a] > 0$
- ii) Für alle $A, \Lambda \in \mathcal{C}(T)$ mit $A \cup \partial A \subset \Lambda$ und alle $b \in \mathcal{U}^{\Lambda \setminus A}$, $a \in \mathcal{U}^A$ ist

$$\mathbf{P}[u_A = a \mid u_{\Lambda \setminus A} = b] = F_{\mathcal{A}}^{b \upharpoonright \partial A}(a)$$

Für ein Potential V bezeichnen wir die **Menge aller Gibbs-Maße mit Potential V** als \mathcal{G}_V .

Eine zweite Art von Maßen auf $(\mathcal{U}^{K_T}, \mathcal{A})$ sind die Markov-Felder.

Definition 11. Ein Wahrscheinlichkeitsmaß $\mathbf{P} \in \mathcal{P}(T)$ heißt **Markov-Feld**, falls gilt

- i) Für alle $A \in \mathcal{C}(T)$ und alle $a \in \mathcal{U}^A$ ist $\mathbf{P}[u_A = a] > 0$.
- ii) Für alle $A, \Lambda \in \mathcal{C}(T)$ mit $A \cup \partial A \subset \Lambda$ und alle $b \in \mathcal{U}^{\Lambda \setminus A}$, $a \in \mathcal{U}^A$ gilt:

$$\mathbf{P}[u_A = a \mid u_{\Lambda \setminus A} = b] = \mathbf{P}[u_A = a \mid u_{\partial \Lambda} = b \upharpoonright \partial \Lambda]$$

Markov-Felder zeichnen sich dadurch aus, daß die Spinkonfiguration einer endlichen Menge nur durch die Spins direkt benachbarter Knoten beeinflusst wird. Dies legt die Vermutung nahe, daß es einen Zusammenhang zu den Gibbs-Maßen gibt, bei denen auch nur direkte Nachbarn zum Potential einer Konfiguration beitragen. Tatsächlich sind Markov-Felder und Gibbs-Maße im wesentlichen gleich, wie folgender Satz zeigt.

Satz 6 ([10], Theorem 4.1). *Sei $\mathbf{P} \in \mathcal{P}(T)$ ein Wahrscheinlichkeitsmaß derart, daß für alle $A \in \mathcal{C}(T)$, $a \in \mathcal{U}^A$ gilt $\mathbf{P}[u_A = a] > 0$. Dann sind folgende Aussagen äquivalent*

- i) \mathbf{P} ist ein Markov-Feld.
- ii) *Es gibt ein eindeutig bestimmtes Nachbarschaftspotential V , sodaß \mathbf{P} ein Gibbs-Maß mit Potential V ist.*

Ein Gibbs-Maß legt das zu ihm gehörige Potential eindeutig fest. Umgekehrt kann es aber zu einem Potential mehrere Gibbs-Maße geben. Dieses Phänomen wird als Phasenübergang bezeichnet. Wir definieren diesen Begriff:

Definition 12. *Sei V ein Potential auf T . Hat \mathcal{G}_V , die Menge aller Gibbs-Maße zu V , mehr als ein Element, so sagen wir, daß **Phasenübergang für V vorliegt**.*

Wir geben nun noch ein Korollar an, das uns dabei helfen wird zu überprüfen, ob ein Markov-Feld für ein gegebenes Potential in \mathcal{G}_V liegt. Das Korollar ist eine Kombination aus Definition 11 und Satz 6.

Korollar 1. *Sei V ein Nachbarschaftspotential auf K_T . $\mathbf{P} \in \mathcal{P}(T)$ sei ein Markov-Feld. Dann sind folgende Aussagen äquivalent*

- i) $\mathbf{P} \in \mathcal{G}_V$
- ii) *Für alle $A, \Lambda \in \mathcal{C}(T)$ mit $A \cup \partial A \subset \Lambda$ und alle $a \in \mathcal{U}^A, b \in \mathcal{U}^{\partial A}$ gilt:*

$$\mathbf{P}[u_A = a \mid u_{\partial A} = b] = \frac{\exp[V((a, b))]}{\sum_{c \in \mathcal{U}^{\partial A}} \exp[V((c, b))]}$$

Dieses Korollar sagt, daß es genügt, die Verteilung auf endlichen Ausschnitten zu untersuchen. Wir werden dies später benutzen, wenn wir zeigen, daß das Broadcasting-Modell eine Darstellung als Gibbs-Maß hat.

Die übliche Art ein Gibbs-Maß zu definieren, verwendet den *thermodynamischen Grenzübergang*. Damit ist gemeint, daß man zu einer Folge endlicher Ausschnitte von T , die T ausschöpfen, ein Folge von Maßen definiert. Das Gibbs-Maß auf T erhält man dann als den Grenzwert dieser Folge von Maßen. Daß diese Vorgehensweise Erfolg hat, besagt folgendes Lemma:

Lemma 1 ([10], Lemma 5.1). *Gegeben sei ein Nachbarschaftspotential V auf T und eine T ausschöpfende Folge endlicher Teilmengen $\Lambda_n \in \mathcal{C}(T)$, kurz: $\Lambda_n \uparrow T$.*

$b_n \in \mathcal{U}^{\partial\Lambda_n}$ sei eine Folge von Randbedingungen. Wir definieren zur Folge $((\Lambda_n), (b_n))$ eine Folge $\mu_n \in \mathcal{P}(T)$ mit:

$$\mu_n(u_{\leq \Lambda_n} = a) := \frac{\exp[V((a, b_n))]}{\sum_{c \in \mathcal{U}^{\leq \Lambda_n}} \exp[V((c, b_n))]} , \quad \text{für } a \in \mathcal{U}^{\leq \Lambda_n}$$

$$\mu_n(u_\sigma = +1) = 1 , \quad \text{falls } \sigma \in \partial\Lambda_n, b_n \rfloor_\sigma = +1$$

$$\mu_n(u_\sigma = -1) = 1 , \quad \text{falls } \sigma \in \partial\Lambda_n, b_n \rfloor_\sigma = -1$$

$$\mu_n(u_\sigma = -1) = 1 , \quad \text{falls } \sigma \notin \Lambda_n \cup \partial\Lambda_n$$

Dann konvergiert die Folge der μ_n gegen ein $\mu \in \mathcal{G}_V$.

Die gleiche Vorgehensweise bringt auch für Gibbs-Maße mit freien Randbedingungen Erfolg. Sei eine Folge von Maßen ψ_n definiert mit:

$$\psi_n(u_{\leq \Lambda_n} = a) := \frac{\exp[V(a)]}{\sum_{c \in \mathcal{U}^{\leq \Lambda_n}} \exp[V(c)]} , \quad \text{falls } a \in \mathcal{U}^{\leq \Lambda_n}$$

$$\psi_n(u_\sigma = -1) = 1 , \quad \text{falls } \sigma \notin \leq \Lambda_n$$

Dann gilt $\psi_n \rightarrow \psi \in \mathcal{G}(V)$.

Wir werden nun zwei ausgezeichnete Maße in der Menge aller Gibbs-Maße zu einem vorgegebenen Potential definieren.

Im folgenden sei V ein fest gewähltes Nachbarschaftspotential auf T . Wir führen folgende Bezeichnungen ein:

$\mathbf{P} \in \mathcal{G}_V$ sei ein Gibbs-Maß zum Potential V . Wir definieren die zu \mathbf{P} gehörige Korrelationsfunktion $\rho : \mathcal{C}(T) \rightarrow [0, 1]$ mit $\rho(A) := \mathbf{P}[u_A \equiv +1]$.

Für $A \subset \Lambda \in \mathcal{C}(T)$ und $x \in \mathcal{U}^{T \setminus \Lambda}$ sei:

$$\rho_\Lambda^x(A) := \mathbf{P}[u_A \equiv +1 \mid u_{T \setminus \Lambda} = x] = \mathbf{P}[u_A \equiv +1 \mid u_{\partial\Lambda} = x]_{\partial\Lambda}$$

und

$$\rho_{\Lambda}^{+}(A) := \max_{x \in \mathcal{U}^{T \setminus \Lambda}} \rho_{\Lambda}^x = \max_{x \in \mathcal{U}^{\partial \Lambda}} \rho_{\Lambda}^x$$

$$\rho_{\Lambda}^{-}(A) := \min_{x \in \mathcal{U}^{T \setminus \Lambda}} \rho_{\Lambda}^x = \min_{x \in \mathcal{U}^{\partial \Lambda}} \rho_{\Lambda}^x$$

Mit diesen Bezeichnungen gilt für alle $A \subset \Lambda \in \mathcal{C}(T)$:

$$\rho_{\Lambda}^{-}(A) \leq \rho(A) \leq \rho_{\Lambda}^{+}(A) \quad (2.12)$$

In [10] wird gezeigt, daß ρ_{Λ}^{-} und ρ_{Λ}^{+} monoton in Λ sind. Für alle $A, \Lambda, \Lambda' \in \mathcal{C}(T)$ mit $A \subset \Lambda \subset \Lambda'$ gilt also:

$$\rho_{\Lambda}^{-}(A) \leq \rho_{\Lambda'}^{-}(A) \quad \text{und} \quad \rho_{\Lambda}^{+}(A) \geq \rho_{\Lambda'}^{+}(A) \quad (2.13)$$

Sei nun Λ_n eine Folge endlicher Teilmengen aus K_T mit $\Lambda_n \uparrow K_T$. Dann existieren nach (2.13) für alle $A \in \mathcal{C}(T)$ folgende Grenzwerte:

$$\rho^{-}(A) := \lim_{\Lambda_n \uparrow K_T} \rho_{\Lambda_n}^{-}(A)$$

$$\rho^{+}(A) := \lim_{\Lambda_n \uparrow K_T} \rho_{\Lambda_n}^{+}(A)$$

In [10] wird weiterhin gezeigt, daß es Maße $\mathbf{P}^{+}, \mathbf{P}^{-} \in \mathcal{G}_V$ gibt, sodaß für alle $A \in \mathcal{C}(T)$ gilt:

$$\mathbf{P}^{+}[u_A \equiv +1] = \rho^{+}(A), \mathbf{P}^{-}[u_A \equiv -1] = \rho^{-}(A)$$

Wir nennen $\mathbf{P}^{+}, \mathbf{P}^{-} \in \mathcal{G}_V$ die **Plus-** bzw. **Minus-Phase** in \mathcal{G}_V .

Aus dem bisher gesagten, insbesondere aus (2.12) folgt, daß Minus- und Plus-Phase die Maße sind, die dem Ereignis u_A die kleinste bzw. größte Wahrscheinlichkeit geben. In diesem Sinne sind sie *extrem*.

Man kann darüber hinaus noch zeigen, daß \mathcal{G}_V konvex und \mathbf{P}^{+} und \mathbf{P}^{-} Extrempunkte in \mathcal{G}_V sind. Wir werden dieses Ergebnis allerdings nicht verwenden.

Wir formulieren eine später verwendete Aussage, die eine direkte Konsequenz der obigen Überlegungen ist, als Satz

Satz 7 ([10]). *Sei V ein Nachbarschaftspotential auf T . \mathbf{P}^{+} sei die Plus-Phase in \mathcal{G}_V , \mathbf{P}^{-} sei die Minus-Phase in \mathcal{G}_V . Dann gilt für alle $\mathbf{P} \in \mathcal{G}_V$ und alle $A \in \mathcal{C}(T)$:*

$$\mathbf{P}^{-}[u_A \equiv +1] \leq \mathbf{P}[u_A \equiv +1] \leq \mathbf{P}^{+}[u_A \equiv +1]$$

Die Aussage aus Satz 7 läßt uns vermuten, daß es einen engen Zusammenhang zwischen dem Auftreten von Phasenübergang und dem Zusammenfallen von Minus- und Plus-Phase gibt. Dies wird im nächsten Satz konkretisiert.

Satz 8 ([10], Theorem 5.7). *Sei V ein nächstes Nachbarschaftspotential auf T . \mathbf{P}^+ und \mathbf{P}^- seien die Plus- und die Minus-Phase in \mathcal{G}_V . Dann gilt:*

$$\text{In } \mathcal{G}_V \text{ tritt Phasenübergang auf} \iff \mathbf{P}^+ \neq \mathbf{P}^-$$

Satz 7 besagt, daß Plus- und Minus-Phase *extreme* Wahrscheinlichkeitsmaße in \mathcal{G}_V sind. Wir werden nun eine Halbordnung auf \mathcal{G}_V einführen, bezüglich der \mathbf{P}^+ und \mathbf{P}^- das maximale bzw. minimale Element in \mathcal{G}_V ist.

Um sagen zu können, was wir unter einer monotonen Abbildung von \mathcal{U}^{K_T} nach \mathbb{R} verstehen wollen, erklären wir zunächst eine Halbordnung auf \mathcal{U}^{K_T} .

Definition 13. *Wir wollen für $a, b \in \mathcal{U}^{K_T}$ sagen $a \leq b$, wenn für alle Knoten, die in a den Spin $+1$ haben, gilt, daß sie auch in b den Spin $+1$ haben. Wir definieren also:*

$$a \leq b :\iff \{\tau \in K_T : a|_\tau = 1\} \subseteq \{\tau \in K_T : b|_\tau = 1\} \quad (2.14)$$

Nun können wir sagen, wann eine Abbildung von \mathcal{U}^{K_T} nach \mathbb{R} monoton ist.

Definition 14. *Auf \mathcal{U}^{K_T} sei die Halbordnung aus Definition 13 erklärt. Wir nennen die Menge $\mathcal{M}(\mathcal{U}^{K_T}, \mathbb{R})$ mit*

$$\mathcal{M}(\mathcal{U}^{K_T}, \mathbb{R}) := \{f : \mathcal{U}^{K_T} \rightarrow \mathbb{R} : \forall a, b \in \mathcal{U}^{K_T}, a \leq b \Rightarrow f(a) \leq f(b)\} \quad (2.15)$$

die Menge der monotonen Funktionen von \mathcal{U}^{K_T} nach \mathbb{R} .

Mit Hilfe dieser Menge können wir nun eine Halbordnung auf \mathcal{G}_V erklären.

Definition 15. \mathbf{P}_1 und \mathbf{P}_2 seien zwei Wahrscheinlichkeitsmaße aus \mathcal{G}_V . Wir schreiben $\mathbf{P}_1 \leq \mathbf{P}_2$, wenn für alle $f \in \mathcal{M}(\mathcal{U}^{K_T}, \mathbb{R})$ gilt

$$\int f d\mathbf{P}_1 \leq \int f d\mathbf{P}_2 \quad (2.16)$$

Der folgende Satz besagt, daß die Plus- und die Minus-Phase bezüglich dieser Halbordnung die maximalen und minimalen Elemente in \mathcal{G}_V sind.

Satz 9 ([6], Korollar 1.30). *Seien \mathbf{P}^+ und \mathbf{P}^- die Plus- und die Minus-Phase in \mathcal{G}_V . Dann gilt mit der Halbordnung aus Definition 15 für alle $\mathbf{P} \in \mathcal{G}_V$:*

$$\mathbf{P}^- \leq \mathbf{P} \leq \mathbf{P}^+ \quad (2.17)$$

Kapitel 3

Dominanz und gegenseitige Information

Wir führen in diesem Kapitel einige Größen ein, die wir zur Untersuchung des Broadcasting-Modells benutzen werden. Alle Begriffe können in einem allgemeineren Rahmen definiert werden, wir beschränken uns jedoch auf die Spezialfälle, wie sie uns im Laufe dieser Arbeit begegnen werden. Für eine weitergehende Betrachtung der gegenseitigen Information zwischen Zufallsvariablen sei auf [1] verwiesen, wo sich auch die hier nicht aufgeführten Beweise befinden.

Im Broadcasting-Modell, welches wir im nächsten Kapitel einführen, wird der Spin der Wurzel eines Baumes T gleichverteilt aus $\{-1, +1\}$ gezogen. Er beeinflusst die Spins der anderen Knoten und so stellt sich die Frage, wie stark die Spinkonfiguration einer Knotenmenge $A \in \mathcal{C}(T)$ vom Spin der Wurzel abhängt. Eine Möglichkeit, diesen Einfluß zu messen, bietet der Begriff der *gegenseitigen Information* zwischen u_r und u_A . Wir führen die gegenseitige Information in Definition 16 für allgemeinere Zufallsvariablen ein.

Sofern nicht anders gesagt, seien in diesem Kapitel X, Y, Z Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum. Das zugrundeliegende Wahrscheinlichkeitsmaß bezeichnen wir mit \mathbf{P} . X, Y, Z mögen nur endlich viele Werte annehmen, ihre Wertebereiche seien \mathcal{X}, \mathcal{Y} bzw. \mathcal{Z} .

Definition 16. Seien X, Y Zufallsvariablen wie oben beschrieben. Dann ist die **gegenseitige Information** zwischen X und Y definiert als:

$$I[X, Y] := \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} \mathbf{P}[X = x, Y = y] \cdot \log \left[\frac{\mathbf{P}[X = x, Y = y]}{\mathbf{P}[X = x] \mathbf{P}[Y = y]} \right]$$

Die gegenseitige Information ist der Kullback-Leibler-Abstand von gemeinsamer Verteilung zur Produktverteilung. Sie ist also ein Abstands begriff für Verteilungen. Wir werden dies noch weiter ausführen und einen weiteren Abstands begriff einführen. Zunächst fassen wir jedoch einige elementare Eigenschaften der gegenseitigen Information in folgendem Lemma zusammen:

Lemma 2. Seien X, Y, Z und Y_1, \dots, Y_n Zufallsvariablen auf dem gleichen Wahrscheinlichkeitsraum. Alle Zufallsvariablen mögen nur endlich viele Werte annehmen und es gebe alle im folgenden auftretenden gemeinsamen Verteilungen. Mit diesen Voraussetzungen gilt:

- (i) $I[X, Y] \geq 0$
- (ii) $I[X, Y] = 0 \iff X, Y \text{ sind stochastisch unabhängig}$
- (iii) Falls $X \rightarrow Y \rightarrow Z$ eine Markov-Kette ist, gilt

$$I[X, Y] \geq I[X, Z]$$

Diese Aussage wird in der Informationstheorie und auch im weiteren Verlauf dieser Arbeit als **Data-processing-Lemma** bezeichnet.

- (iv) Falls Y_1, \dots, Y_n bedingt auf X unabhängig sind, so gilt für die Information zwischen X und dem Vektor (Y_1, \dots, Y_n) :

$$I[X, (Y_1, \dots, Y_n)] \leq \sum_{i=1}^n I[X, Y_i]$$

Wir führen nun einen Dominanz-Begriff für Paare von Zufallsvariablen ein, der eine obere Abschätzung für die gegenseitige Information liefert.

Definition 17. Wir sagen, das Paar (X, Y) **dominiert** das Paar (X, Z) , wenn es eine Übergangsmatrix $D = (d(y, z))_{\substack{y \in \mathcal{Y} \\ z \in \mathcal{Z}}}$ gibt, sodaß für alle $x \in \mathcal{X}$ und $z \in \mathcal{Z}$ gilt

$$\mathbf{P}[Z = z | X = x] = \sum_{y \in \mathcal{Y}} d(y, z) \cdot \mathbf{P}[Y = y | X = x] \quad (3.1)$$

Dieser Dominanzbegriff orientiert sich an dem in [11] eingeführten Dominanzbegriff für Bäume.

Der nächste Satz zeigt, daß eine enge Beziehung zwischen der Dominanz eines Paares und der Existenz einer Markov-Kette besteht. Diese Tatsache werden wir uns bald zunutze machen, um die Information zwischen X und Y mit der Information zwischen X und Z zu vergleichen:

Satz 10. *Mit den Bezeichnungen von oben gilt:*

(X, Y) dominiert (X, Z)

$$\iff \text{Es gibt eine Markov-Kette } X \rightarrow Y \rightarrow \tilde{Z} \text{ mit } \mathcal{L}(X, Z) = \mathcal{L}(X, \tilde{Z})$$

Beweis.

- i) Wenn das Paar (X, Y) das Paar (X, Z) dominiert, definieren wir die gemeinsame Verteilung von X, Y, \tilde{Z} als:

$$\mathbf{P}[X = x, Y = y, \tilde{Z} = z] = \mathbf{P}[X = x, Y = y] \cdot d(y, z) \quad (3.2)$$

Damit gilt

$$\begin{aligned} \mathbf{P}[\tilde{Z} = z | Y = y, X = x] &= \frac{\mathbf{P}[\tilde{Z} = z, Y = y, X = x]}{\mathbf{P}[Y = y, X = x]} \\ &= \frac{\mathbf{P}[Y = y, X = x]}{\mathbf{P}[Y = y, X = x]} \cdot d(y, z) \\ &= d(y, z) \\ &= \mathbf{P}[\tilde{Z} = z | Y = y] \end{aligned}$$

Demnach ist $X \rightarrow Y \rightarrow \tilde{Z}$ eine Markov-Kette. Weiterhin stimmen auch die Verteilungen von (X, Z) und (X, \tilde{Z}) überein, denn für alle $x \in \mathcal{X}$, $z \in \mathcal{Z}$ gilt:

$$\begin{aligned} \mathbf{P}[X = x, Z = z] &= \sum_{y \in \mathcal{Y}} d(y, z) \cdot \mathbf{P}[Y = y | X = x] \cdot \mathbf{P}[X = x] \\ &= \mathbf{P}[X = x, \tilde{Z} = z] \end{aligned}$$

- ii) Gibt es andererseits eine Markov-Kette $X \rightarrow Y \rightarrow \tilde{Z}$, sodaß das Paar (X, Z) so verteilt ist wie das Paar (X, \tilde{Z}) , so definieren wir die Einträge der Übergangsmatrix $D = (d(y, z))_{\substack{y \in \mathcal{Y} \\ z \in \mathcal{Z}}}$ als:

$$d(y, z) := \mathbf{P}[\tilde{Z} = z \mid Y = y]$$

Dann folgt sofort:

$$\begin{aligned} \mathbf{P}[Z = z \mid X = x] &= \mathbf{P}[\tilde{Z} = z \mid X = x] \\ &= \sum_{y \in \mathcal{Y}} \mathbf{P}[\tilde{Z} = z \mid Y = y, X = x] \cdot \mathbf{P}[Y = y \mid X = x] \\ &= \sum_{y \in \mathcal{Y}} \mathbf{P}[\tilde{Z} = z \mid Y = y] \cdot \mathbf{P}[Y = y \mid X = x] \\ &= \sum_{y \in \mathcal{Y}} d(y, z) \cdot \mathbf{P}[Y = y \mid X = x] \end{aligned}$$

Also wird (X, Z) von (X, Y) dominiert.

□

Als eine einfache Folgerung aus diesem Satz erhalten wir ein Korollar, das die Dominanz mit der gegenseitigen Information in Verbindung bringt.

Korollar 2. *Wenn das Paar (X, Y) das Paar (X, Z) dominiert, gilt*

$$I[X, Y] \geq I[X, Z]$$

Beweis. Nach Satz 10 folgt aus der Dominanz von (X, Y) über (X, Z) , daß es eine Markov-Kette $X \rightarrow Y \rightarrow \tilde{Z}$ gibt mit $\mathcal{L}(X, Z) = \mathcal{L}(X, \tilde{Z})$. Damit ist auch $I[X, Z] = I[X, \tilde{Z}]$. Das Data-processing-Lemma besagt nun

$$I[X, Y] \geq I[X, \tilde{Z}] = I[X, Z]$$

□

Wir beschreiben nun das Szenario, in dem uns die gegenseitige Information beschäftigen wird, etwas genauer und führen einen zweiten Abstandsbegriff ein.

Sei \mathbf{P} ein Wahrscheinlichkeitsmaß auf $(\mathcal{U}^{K_T}, \mathcal{A})$, welches die Eigenschaft erfüllt:

$$\mathbf{P}[u_r = +1] = \mathbf{P}[u_r = -1] = \frac{1}{2} \quad (3.3)$$

Wir fragen uns nun, ob wir anhand der Beobachtung einer Größe $S = S(u_A)$, die von der Spinkonfiguration auf A abhängt, auf den Spin der Wurzel schließen können. Dies wird nur dann möglich sein, wenn sich die bedingte Verteilung von u_A gegeben $u_r = -1$ von der bedingten Verteilung gegeben $u_r = +1$ unterscheidet. Um dies zu untersuchen, benutzen wir zwei Abstandsbegriffe, deren Definition einiger Bezeichnungen bedarf.

$A \in \mathcal{C}(T)$ sei eine fest gewählte Menge von Knoten. S sei eine Abbildung von \mathcal{U}^A in eine endliche Menge $\mathcal{X} = S(\mathcal{U}^A)$, der Menge der Beobachtungen. In den späteren Anwendungen ist S zum einen die Spinkonfiguration selbst, zum anderen eine gewichtete Summe der Spins in A . Für $x \in \mathcal{X}$ sei:

$$\nu(x) := \mathbf{P}[u_A \in S^{-1}(x)] \quad (3.4)$$

$$\nu_-(x) := \mathbf{P}[u_A \in S^{-1}(x) \mid u_r = -1] \quad (3.5)$$

$$\nu_+(x) := \mathbf{P}[u_A \in S^{-1}(x) \mid u_r = +1] \quad (3.6)$$

Wegen (3.3) gilt für alle $x \in \mathcal{X}$:

$$\nu(x) = \frac{1}{2} \cdot \nu_-(x) + \frac{1}{2} \cdot \nu_+(x)$$

Weiterhin seien $f, g : \mathcal{X} \rightarrow [0, 2]$ definiert als:

$$\begin{aligned} f(x) &:= 2 \cdot \mathbf{P}[u_r = +1 \mid u_A \in S^{-1}(x)] \\ &= \frac{\nu_+(x)}{\nu(x)} \\ &= \frac{\nu_+(x)}{\frac{1}{2}\nu_+(x) + \frac{1}{2}\nu_-(x)} \end{aligned}$$

$$\begin{aligned}
g(x) &:= 2 \cdot \mathbf{P}[u_r = -1 \mid u_A \in S^{-1}(x)] \\
&= \frac{\nu_-(x)}{\nu(x)} \\
&= \frac{\nu_-(x)}{\frac{1}{2}\nu_+(x) + \frac{1}{2}\nu_-(x)}
\end{aligned}$$

Die gegenseitige Information zwischen u_r und $S(u_A)$ liefert die erste Möglichkeit, den Abstand zwischen ν_+ und ν_- zu messen:

$$D_I(\nu_+, \nu_-) := I[u_r, S(u_A)] \quad (3.7)$$

$$\begin{aligned}
&= \frac{1}{2} \cdot \sum_{x \in \mathcal{X}} \left[\frac{\nu_+(x)}{\nu(x)} \cdot \log \left[\frac{\nu_+(x)}{\nu(x)} \right] + \frac{\nu_-(x)}{\nu(x)} \cdot \log \left[\frac{\nu_-(x)}{\nu(x)} \right] \right] \cdot \nu(x) \\
&= \frac{1}{2} \sum_{x \in \mathcal{X}} [f(x) \log[f(x)] + g(x) \log[g(x)]] \cdot \nu(x) \quad (3.8)
\end{aligned}$$

Dieser Abstand ist eine symmetrisierte Version des Kullback-Leibler-Abstands von ν_- und ν_+ , oder, wie bereits erwähnt, die relative Entropie der gemeinsamen Verteilung von u_r und $S(u_A)$ zur Produktverteilung.

Eine andere Möglichkeit, den Abstand zwischen ν_+ und ν_- zu messen, liefert der χ^2 -Abstand. Wir definieren ebenfalls eine symmetrische Variante des üblichen Begriffs als:

$$D_\chi(\nu_+, \nu_-) := \frac{1}{2} \left(\sum_{x \in \mathcal{X}} \frac{(\nu_+(x) - \nu_-(x))^2}{\nu(x)} \right)^{\frac{1}{2}} \quad (3.9)$$

$$= \frac{1}{2} \left(\sum_{x \in \mathcal{X}} (f(x) - g(x))^2 \cdot \nu(x) \right)^{\frac{1}{2}} \quad (3.10)$$

Wir fassen in folgendem Lemma einige Eigenschaften dieser Abstände zusammen:

Lemma 3. *Mit den Bezeichnungen von oben gilt:*

$$\frac{1}{2} (D_\chi(\nu_+, \nu_-))^2 \leq D_I(\nu_+, \nu_-) \leq (D_\chi(\nu_+, \nu_-))^2 \quad (3.11)$$

Ist weiterhin $S : \mathcal{U}^A \rightarrow \mathbb{R}$, so gilt:

$$\left(\sum_{x \in \mathcal{X}} x \cdot (\nu_+(x) - \nu_-(x)) \right)^2 = \left(\sum_{x \in \mathcal{X}} x \cdot (f(x) - g(x)) \cdot \nu(x) \right)^2 \quad (3.12)$$

$$\leq 4 \left(\sum_{x \in \mathcal{X}} x^2 \nu(x) \right) \cdot (D_\chi(\nu_+, \nu_-))^2 \quad (3.13)$$

Beweis. Wir werden eine allgemeinere Form der obigen Aussagen beweisen, da dies aus technischer Sicht einfacher ist. Daher führen wir hier kurz eine allgemeinere Version der oben bereits definierten Größen ein. Nach Beendigung dieses Beweises werden wir uns aber ausschließlich auf die Begriffe beziehen, wie sie oben (in der diskreten Version) eingeführt wurden.

Sei nun also \mathcal{X} eine nicht mehr notwendigerweise endliche Menge und seien ν_+ und ν_- zwei Verteilungen einer Zufallsvariable, die Werte in \mathcal{X} annimmt. Weiterhin definieren wir:

$$\nu := \frac{\nu_- + \nu_+}{2}, \quad f := \frac{d\nu_+}{d\nu}, \quad g := \frac{d\nu_-}{d\nu}$$

Damit ist insbesondere

$$f + g \equiv 2 \quad (3.14)$$

Wir definieren in diesem allgemeineren Szenario $D_I(\nu_+, \nu_-)$ analog wie oben als:

$$D_I(\nu_+, \nu_-) := \frac{1}{2} \int_{\mathcal{X}} f(x) \log[f(x)] + g(x) \log[g(x)] d\nu \quad (3.15)$$

Ebenso D_χ als:

$$D_\chi(\nu_+, \nu_-) := \frac{1}{2} \left(\int_{\mathcal{X}} (f(x) - g(x))^2 d\nu \right)^{\frac{1}{2}} \quad (3.16)$$

Damit gilt:

$$(D_\chi(\nu_+, \nu_-))^2 = \int_{\mathcal{X}} \left(\frac{f(x) - g(x)}{2} \right)^2 d\nu \quad (3.17)$$

Die Aussage in (3.13) ist nun ein Spezialfall der folgenden Aussage

$$\left(\int_{\mathcal{X}} x(d\nu_+ - d\nu_-) \right)^2 = \left(\int_{\mathcal{X}} x(f(x) - g(x)) d\nu \right)^2 \quad (3.18)$$

$$\begin{aligned} &\leq 4 \cdot \int_{\mathcal{X}} x^2 d\nu \cdot (D_\chi(\nu_+, \nu_-))^2 \\ &= \int_{\mathcal{X}} x^2 d\nu \cdot \int_{\mathcal{X}} (f(x) - g(x))^2 d\nu \end{aligned} \quad (3.19)$$

Die Aussage aus Gleichung (3.18) gilt nach Definition von f und g . Ein Vergleich der rechten Seiten von (3.18) und (3.19) zeigt, daß die Richtigkeit dieser Aussage durch die Cauchy-Schwarz'sche Ungleichung garantiert ist.

Wenden wir uns nun den Ungleichungen aus (3.11) zu. Wir führen eine Abbildung $\psi : \mathcal{X} \mapsto (-1, 1)$ ein mit:

$$\psi := \frac{f - g}{2} \quad (3.20)$$

Damit erhalten wir für D_χ^2 aus (3.17) die Darstellung

$$(D_\chi(\nu_+, \nu_-))^2 = \int_{\mathcal{X}} \psi(x)^2 d\nu \quad (3.21)$$

Unter Verwendung von (3.14) und (3.20) läßt sich auch D_I anders schreiben:

$$D_I(\nu_+, \nu_-) = \int_{\mathcal{X}} \frac{1 + \psi(x)}{2} \log[1 + \psi(x)] + \frac{1 - \psi(x)}{2} \log[1 - \psi(x)] d\nu \quad (3.22)$$

Damit nimmt die allgemeine Version von Gleichung (3.11) folgende Gestalt an:

$$\frac{1}{2} \int_{\mathcal{X}} \psi^2 d\nu \leq \int_{\mathcal{X}} \frac{1 + \psi}{2} \log[1 + \psi] + \frac{1 - \psi}{2} \log[1 - \psi] d\nu \leq \int_{\mathcal{X}} \psi^2 d\nu \quad (3.23)$$

Wir werden die Richtigkeit dieser Aussage nachweisen, indem wir zeigen, daß die Integranden die entsprechenden Abschätzung punktweise erfüllen.

Für $y \in (-1, 1)$ gilt bekanntlich $\log[1 + y] \leq y$. Damit erhalten wir für alle $x \in \mathcal{X}$:

$$\begin{aligned} & \frac{1 + \psi(x)}{2} \log[1 + \psi(x)] + \frac{1 - \psi(x)}{2} \log[1 - \psi(x)] \\ & \leq \frac{1 + \psi(x)}{2} \cdot \psi(x) + \frac{1 - \psi(x)}{2} \cdot (-\psi(x)) \\ & = \psi(x)^2 \end{aligned}$$

Diese Eigenschaft überträgt sich auf die Integrale und die rechte Ungleichung aus (3.23) ist somit bewiesen.

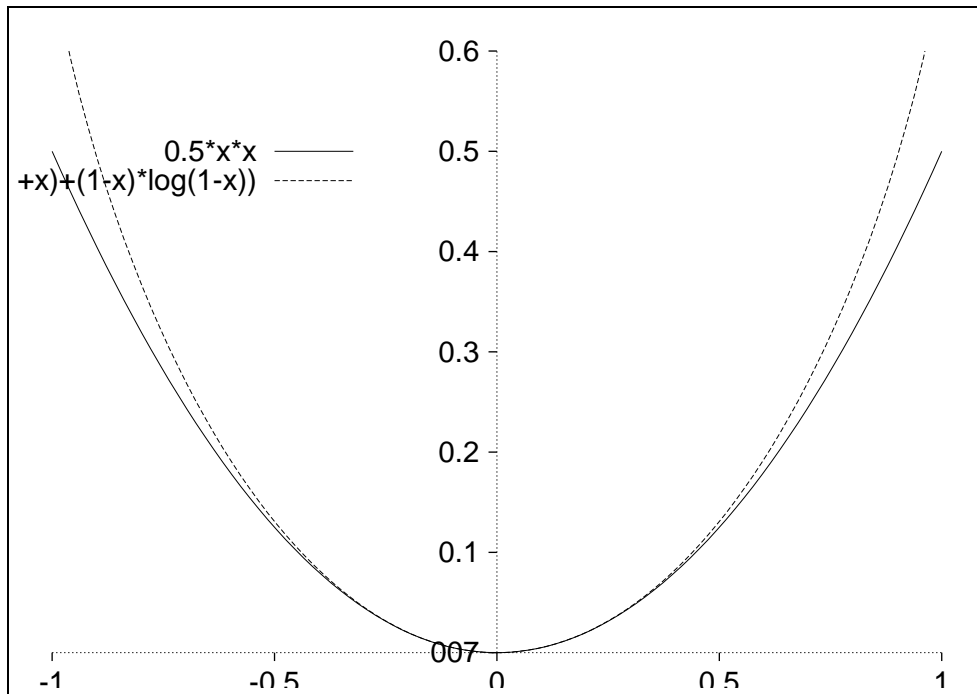


Abbildung 3.1: $h(z)$ und $i(z)$ aus dem Beweis der linken Ungleichung aus (3.23)

Um die Richtigkeit der linken Ungleichung zu erkennen, substituieren wir

$z := \psi(x)$ und betrachten die Abbildungen:

$$h : (-1, 1) \rightarrow \mathbb{R}, \quad z \mapsto h(z) := \frac{z^2}{2}$$

$$i : (-1, 1) \rightarrow \mathbb{R}, \quad z \mapsto i(z) := \frac{1+z}{2} \log[1+z] + \frac{1-z}{2} \log[1-z]$$

Einerseits ist $h(0) = 0 = i(0)$, andererseits gilt für die zweiten Ableitungen im Intervall $(-1, 1)$:

$$h''(z) = 1$$

$$i''(z) = \frac{1}{2} \frac{1}{1+z} + \frac{1}{2} \frac{1}{1-z} > 1$$

Damit ist für $z \in (-1, 1)$ stets: $h(z) \leq i(z)$, und dies ist nichts anderes als:

$$\frac{\psi(x)^2}{2} \leq \frac{1+\psi(x)}{2} \log[1+\psi(x)] + \frac{1-\psi(x)}{2} \log[1-\psi(x)]$$

Diese Eigenschaft gilt punktweise für $x \in (-1, 1)$ und überträgt sich somit auf die Integrale in der linken Ungleichung aus (3.23). \square

Kapitel 4

Das Broadcasting-Modell

Wir führen in diesem Kapitel ein Modell ein, das das eigentliche Objekt unserer Untersuchung ist und nennen mehrere Sätze, die sein Verhalten charakterisieren. Einige längere Beweise haben wir in das nächste Kapitel verbannt.

4.1 Grundlegende Eigenschaften des Modells

Das folgende stochastische Modell wird in [11] eingeführt. Es gibt verschiedene anschauliche Interpretationen des Modells, die wir im folgenden erläutern werden.

Definition 18. T sei ein lokal endlicher Baum, der nur unendliche Äste hat, und ϵ sei eine als **Fehlerwahrscheinlichkeit** bezeichnete reelle Zahl mit $0 < \epsilon \leq \frac{1}{2}$. Jeder Kante $e \in E_T$ wird eine $\{+1, -1\}$ -wertige Zufallsvariable η_e zugeordnet. Die η_e seien stochastisch unabhängig und es gelte für alle $e \in E_T$:

$$\mathbf{P}[\eta_e = -1] = \epsilon \quad , \quad \mathbf{P}[\eta_e = 1] = 1 - \epsilon$$

Jedem Knoten $\sigma \in K_T$ wird eine als **Spin** bezeichnete Zufallsvariable u_σ mit Werten in $\{-1, +1\}$ zugeordnet. Der Spin der Wurzel sei dabei uniform verteilt. Für jeden Knoten $\sigma \in K_T, \sigma \neq r$ gelte weiterhin:

$$u_\sigma = u_r \prod_{e \in \text{Pfad}(r, \sigma)} \eta_e \tag{4.1}$$

\mathcal{A} sei die von den endlichdimensionalen Zylindermengen erzeugte σ -Algebra auf \mathcal{U}^{K_T} . $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ heißt dann **Broadcasting-Modell mit Fehlerwahrscheinlichkeit ϵ** .

Eine anschauliche Interpretation des Broadcasting-Modells ist die der Signalübertragung. Dabei wird ein Signal, das $+1$ oder -1 sein kann, zufällig in die Wurzel des Baumes gesetzt und entlang der Äste des Baumes von Knoten zu Knoten weitergegeben. Die Kanten agieren als Übertragungskanäle zwischen den Knoten. Jede Kante kann das Signal korrekt übermitteln (dann, wenn der Wert der ihr zugeordneten Zufallsvariable 1 ist), oder sie kann das Signal flippen, also das Vorzeichen umkehren. Korrekte Übertragung geschieht an allen Kanten unabhängig mit der gleichen Wahrscheinlichkeit $1 - \epsilon$.

Abbildung 4.1 zeigt eine Realisierung des Modells. Dabei steht “ \star ” genau an den Kanten, an denen die Zufallsvariable der Kante -1 ist.

Wie wir später sehen werden, gibt es noch eine andere Möglichkeit dieses Modell zu beschreiben, die ebenfalls eine anschauliche Interpretation hat. Zunächst untersuchen wir jedoch diese Darstellung.

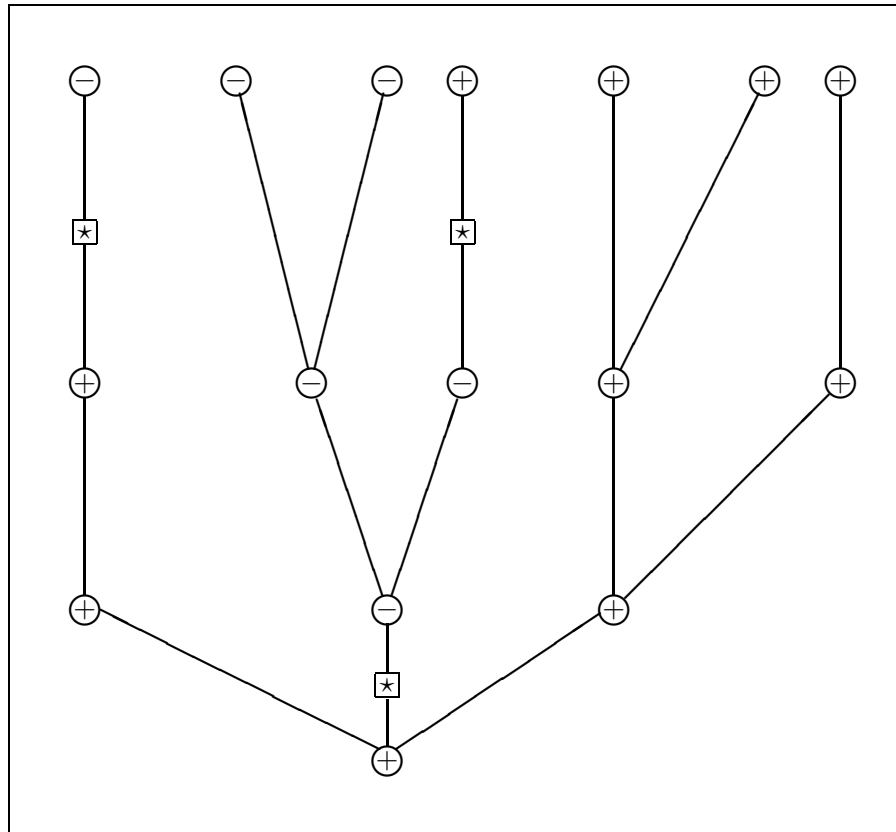


Abbildung 4.1: Eine Realisierung des Broadcasting-Modells

Die Verteilung der u_σ ist natürlich durch die Verteilung der η_e festgelegt. Allerdings ist die Verteilung der Spins nicht so einfach wie die der Zufallsvariablen an den Kanten, insbesondere sind die Spins nicht unabhängig.

Für den Spin der Wurzel gilt nach Konstruktion bereits, daß die Wahrscheinlichkeiten für $+1$ und -1 gleich sind. Dies gilt allerdings auch für alle anderen Knoten, wie folgende kurze Überlegung zeigt:

Sei $\sigma \in K_T$ ein Nachfolgeknoten von der Wurzel r . Dann gilt:

$$\begin{aligned} \mathbf{P}[u_\sigma = 1] &= (1 - \epsilon)\mathbf{P}[u_r = 1] + \epsilon\mathbf{P}[u_r = -1] \\ &= (1 - \epsilon) \cdot \frac{1}{2} + \epsilon \cdot \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

Ein einfaches Induktionsargument liefert sofort, daß für alle $\sigma \in K_T$ gilt:

$$\mathbf{P}[u_\sigma = 1] = \mathbf{P}[u_\sigma = -1] = \frac{1}{2} \quad (4.2)$$

Das Broadcasting-Modell ist ein attraktives Modell in dem Sinne, daß die Wahrscheinlichkeit für gleichen Spin in benachbarten Knoten größer oder gleich der Wahrscheinlichkeit ist, daß die Spins unterschiedlich sind. Gleichen Spin haben die Knoten genau dann, wenn die Kante zwischen ihnen korrekt übermittelt, was mit Wahrscheinlichkeit $1 - \epsilon$ geschieht. Diese ist größer oder gleich der Wahrscheinlichkeit für eine fehlerhafte Kante.

Die Fehlerwahrscheinlichkeit ϵ ist ein Maß dafür, wie stark benachbarte Knoten interagieren. Um dies ein wenig klarer zu machen, betrachten wir die zwei Extremfälle $\epsilon = 0$ und $\epsilon = \frac{1}{2}$. Zwar wird der Fall $\epsilon = 0$ in der Definition 18 nicht erfaßt, aber wir können ihn als Grenzfall von $\epsilon \rightarrow 0$ betrachten.

Für $\epsilon = 0$ gilt für alle $\sigma \in K_T, \sigma \neq r$:

$$\mathbf{P}[u_\sigma = 1 \mid u_{\overline{\sigma}} = 1] = \mathbf{P}[u_\sigma = -1 \mid u_{\overline{\sigma}} = -1] = \mathbf{P}[\eta_{e_{\sigma, \overline{\sigma}}} = 1] = 1$$

Jeder Knoten hat **P**-fast-sicher den gleichen Spin wie sein Vorgänger. Der Spin der Wurzel ist gleichverteilt auf $\{-1, +1\}$, also haben im Fall $\epsilon = 0$ nur zwei Konfigurationen auf T positive Wahrscheinlichkeit, nämlich die Konfiguration, in der alle Knoten Spin $+1$ haben und die, in der alle Knoten Spin -1 haben.

Beide Konfigurationen haben unter \mathbf{P} die Wahrscheinlichkeit $\frac{1}{2}$. Der Fall $\epsilon = 0$ entspricht *perfekter* oder auch *vollständiger Interaktion*. Solchen Interaktionen werden wir im weiteren Verlauf allerdings nicht mehr begegnen.

Der Fall $\epsilon = \frac{1}{2}$ entspricht, wie wir bereits erwarten, dem Fall *keiner Interaktion* oder der *Entkoppelung*. Die Spins benachbarter Knoten sind stochastisch unabhängig, denn es gilt für alle Knoten $\sigma \in K_T, \sigma \neq r$:

$$\mathbf{P}[u_\sigma = 1 \mid u_{\overline{\sigma}} = 1] = \mathbf{P}[u_\sigma = -1 \mid u_{\overline{\sigma}} = -1] = \mathbf{P}[\eta_{e(\sigma, \overline{\sigma})} = 1] = \frac{1}{2}$$

$$\mathbf{P}[u_\sigma = -1 \mid u_{\overline{\sigma}} = 1] = \mathbf{P}[u_\sigma = 1 \mid u_{\overline{\sigma}} = 1] = \mathbf{P}[\eta_{e(\sigma, \overline{\sigma})} = -1] = \frac{1}{2}$$

In diesem Fall ist der Zustand eines Knotens stochastisch unabhängig vom Zustand seines Vorgängers.

Unter Berücksichtigung von $\mathbf{P}[u_r = 1] = \mathbf{P}[u_r = -1] = \frac{1}{2}$ folgt, daß hier \mathbf{P} die Verteilung auf \mathcal{U}^{K_T} ist, die jeder Konfiguration das gleiche Gewicht gibt.

Natürlich sind diese beiden Extremfälle nur von geringem Interesse. Daher wollen wir uns insbesondere mit den Fällen $0 < \epsilon < \frac{1}{2}$ beschäftigen. Um das Broadcasting-Modell in einen allgemeineren Rahmen zu setzen und somit auch allgemeine Sätze zu Rate ziehen zu können, zeigen wir zunächst, daß es ein Markov-Feld ist.

Eine erste Einsicht in die Verteilung der $u_\sigma, \sigma \in K_T$ erhalten wir durch die Betrachtung des bedingten Ereignisses $[u_\sigma = x \mid u_{N(\sigma)} = y]$, wobei $x \in \{-1, +1\}$ und $y \in \mathcal{U}^{N(\sigma)}$ ist.

m sei die Gesamtzahl der Nachbarn von σ , $l(x, y)$ sei die Zahl der Nachbarn von σ , die gleichen Spin haben wie σ und $o(x, y)$ sei die Anzahl der Nachbarn mit entgegengesetztem Spin:

$$l(x, y) := \#\{i \mid y_i = x\} \tag{4.3}$$

$$o(x, y) := \#\{i \mid y_i \neq x\} \tag{4.4}$$

Weiterhin seien die Nachbarn von σ numeriert: $N(\sigma) = \{\tau_1, \dots, \tau_m\}$.

Mit diesen Bezeichnungen gilt:

$$\begin{aligned} \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = x] &= \frac{\mathbf{P}[u_\sigma = x, u_{\tau_1} = y_1, \dots, u_{\tau_m} = y_m]}{\mathbf{P}[u_\sigma = x]} \\ &= \frac{\frac{1}{2} \mathbf{P}[\eta_{e(\sigma, \tau_1)} = xy_1, \dots, \eta_{e(\sigma, \tau_m)} = xy_m]}{\frac{1}{2}} \end{aligned} \quad (4.5)$$

$$\begin{aligned} &= \mathbf{P}[\eta_{e(\sigma, \tau_1)} = xy_1] \cdots \mathbf{P}[\eta_{e(\sigma, \tau_m)} = xy_m] \quad (4.6) \\ &= (1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)} \end{aligned}$$

Die Umformung von (4.5) nach (4.6) ist wegen der Unabhängigkeit der Übertragung an den Kanten möglich. Mit einer analogen Rechnung erhalten wir sofort:

$$\mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = -x] = (1 - \epsilon)^{o(x, y)} \epsilon^{l(x, y)} \quad (4.7)$$

Unter Verwendung des Satzes von Bayes erhalten wir nun die gesuchte Wahrscheinlichkeit:

$$\begin{aligned} \mathbf{P}[u_\sigma = x \mid u_{N(\sigma)} = y] &= \frac{\frac{1}{2} \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = x]}{\frac{1}{2} \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = x] + \frac{1}{2} \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = -x]} \\ &= \frac{(1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)}}{(1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)} + (1 - \epsilon)^{o(x, y)} \epsilon^{l(x, y)}} \end{aligned} \quad (4.8)$$

Wie wir bald sehen werden, wird das Broadcasting-Modell durch Gleichung (4.8) charakterisiert.

Der folgende Satz wird das Broadcasting-Modell als ein Markov-Feld enttarnen.

Satz und Definition 2. *Das Broadcasting-Modell ist ein Markov-Feld. Dabei gilt für alle Knoten $\sigma \in K_T$, mit $x \in \{-1, +1\}$, $a \in \mathcal{U}^{K_T \setminus \sigma}$ und $y := a|_{N(\sigma)}$ die lokale Charakteristik:*

$$\begin{aligned} \mathbf{P}[u_\sigma = x \mid u_{K_T \setminus \sigma} = a] &= \mathbf{P}[u_\sigma = x \mid u_{N(\sigma)} = y] \\ &= \frac{(1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)}}{(1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)} + (1 - \epsilon)^{o(x, y)} \epsilon^{l(x, y)}} \end{aligned} \quad (4.9)$$

Beweis. Sei $\sigma \in K_T$, $x \in \{-1, +1\}$ und $a \in \{-1, +1\}^{K_T \setminus \sigma}$. Weiterhin sei $a' = a|_{K_T \setminus N(\sigma)}$ und $y = a|_{N(\sigma)}$.

Eine Belegung von $\hat{N}(\sigma)$ schreiben wir als $u_{\hat{N}(\sigma)} = (x, y)$, wobei dies so zu lesen ist, daß $u_\sigma = x$ und $u_{N(\sigma)} = y$ ist. Mit diesen Bezeichnungen gilt:

$$\begin{aligned}
& \mathbf{P}[u_\sigma = x \mid u_{K_T \setminus \sigma} = a] = \\
&= \frac{\mathbf{P}[u_\sigma = x] \mathbf{P}[u_{K_T \setminus \sigma} = a \mid u_\sigma = x]}{\mathbf{P}[u_\sigma = -x] \mathbf{P}[u_{K_T \setminus \sigma} = a \mid u_\sigma = -x] + \mathbf{P}[u_\sigma = x] \mathbf{P}[u_{K_T \setminus \sigma} = a \mid u_\sigma = x]} \\
&= \frac{\mathbf{P}[u_{K_T \setminus \sigma} = a \mid u_\sigma = x]}{\mathbf{P}[u_{K_T \setminus \sigma} = a \mid u_\sigma = -x] + \mathbf{P}[u_{K_T \setminus \sigma} = a \mid u_\sigma = x]} \\
&= \frac{\mathbf{P}[u_{K_T \setminus \hat{N}(\sigma)} = a', u_{N(\sigma)} = y, u_\sigma = x]}{\mathbf{P}[u_{K_T \setminus \hat{N}(\sigma)} = a', u_{N(\sigma)} = y, u_\sigma = -x] + \mathbf{P}[u_{K_T \setminus \hat{N}(\sigma)} = a', u_{N(\sigma)} = y, u_\sigma = x]} \\
&= \frac{\mathbf{P}[u_{K_T \setminus \hat{N}(\sigma)} = a', u_{N(\sigma)} = y]}{\mathbf{P}[u_{K_T \setminus \hat{N}(\sigma)} = a', u_{N(\sigma)} = y]} \cdot \frac{\mathbf{P}[u_{\hat{N}(\sigma)} = (x, y)]}{\mathbf{P}[u_{\hat{N}(\sigma)} = (-x, y)] + \mathbf{P}[u_{\hat{N}(\sigma)} = (x, y)]} \\
&= \frac{\mathbf{P}[u_{N(\sigma)} = y, u_\sigma = x]}{\mathbf{P}[u_{N(\sigma)} = y, u_\sigma = -x] + \mathbf{P}[u_{N(\sigma)} = y, u_\sigma = x]} \\
&= \frac{\mathbf{P}[u_\sigma = x] \cdot \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = x]}{\mathbf{P}[u_\sigma = x] \cdot \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = x] + \mathbf{P}[u_\sigma = -x] \cdot \mathbf{P}[u_{N(\sigma)} = y \mid u_\sigma = -x]} \\
&= \mathbf{P}[u_\sigma = x \mid u_{N(\sigma)} = y] \\
&= \frac{(1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)}}{(1 - \epsilon)^{l(x, y)} \epsilon^{o(x, y)} + (1 - \epsilon)^{o(x, y)} \epsilon^{l(x, y)}}
\end{aligned}$$

Wir haben in dieser Umformung verwendet, daß die Ereignisse

$$(u_{K_T \setminus \hat{N}(\sigma)} = a', u_{N(\sigma)} = y) \quad \text{und} \quad (u_{N(\sigma)} = y, u_\sigma = x)$$

stochastisch unabhängig sind. Im linken Ereignis sind nur Kanten involviert, die außerhalb von $\hat{N}(\sigma)$ liegen, im rechten Ereignis tauchen nur Kanten auf, die zwischen σ und seinen Nachbarn liegen. Da das Übertragungsverhalten der Kanten unabhängig ist, sind es auch diese Ereignisse. \square

Die Erkenntnis, daß das Broadcasting-Modell die Markov-Feld-Eigenschaft erfüllt, wird uns bei vielen Rechnungen hilfreich sein. So auch in den Beweisen des nächsten Satzes, der einige Eigenschaften des Broadcasting-Modells zusammenfaßt.

Satz 11. *Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Broadcasting-Modell mit Fehlerwahrscheinlichkeit ϵ . σ sei ein beliebiger Knoten, T^n sei die n -Sphäre. Dann gilt:*

- (i) $\mathbf{E}[u_r \cdot u_\sigma] = (1 - 2\epsilon)^{|\sigma|} =: \theta^{|\sigma|}$
- (ii) $\mathbf{P}[u_{T^n} \equiv -1 \mid u_r = +1] \leq \mathbf{P}[u_{T^n} \equiv +1 \mid u_r = +1] \leq (1 - \epsilon)^{\#T^n}$
- (iii) $\lim_{n \rightarrow \infty} \mathbf{P}[u_{T^n} \equiv +1] = \lim_{n \rightarrow \infty} \mathbf{P}[u_{T^n} \equiv +1 \mid u_r = +1] = 0$
- (iv) *Mit der in Definition 16 eingeführten gegenseitigen Information gilt:*

$$I[u_r, u_\sigma] \leq (\mathbf{E}[u_r u_\sigma])^2 = \theta^{2|\sigma|} \quad (4.10)$$

Beweis.

(i)

$$\begin{aligned} \mathbf{E}[u_r \cdot u_\sigma] &= \mathbf{E} \left[u_r \cdot u_\sigma \cdot \prod_{e \in \text{Pfad}(r, \sigma)} \eta_e \right] \\ &= \mathbf{E}[u_r^2] \prod_{e \in \text{Pfad}(r, \sigma)} \mathbf{E}[\eta_e] \\ &= (1 - 2\epsilon)^{|\sigma|} \end{aligned}$$

- (ii) Die Attraktivität des Broadcasting-Modells liefert die erste Ungleichung. Die zweite Ungleichung gilt, denn:

$$\begin{aligned} \mathbf{P}[u_{T^n} \equiv +1 \mid u_r = +1] &\leq \mathbf{P}[u_{T^n} \equiv +1 \mid u_{T^{n-1}} \equiv +1, u_r = +1] \\ &= \mathbf{P}[u_{T^n} \equiv +1 \mid u_{T^{n-1}} \equiv +1] \\ &= (1 - \epsilon)^{\#T^n} \end{aligned}$$

(iii) Wir beschäftigen uns nur mit Bäumen mit $\text{brT} > 1$. Daher gilt

$$\lim_{n \rightarrow \infty} \#T^n = \infty$$

Unter Verwendung der Symmetrie des Broadcasting-Modells folgt mit der ersten Ungleichung aus (ii):

$$\mathbf{P}[u_{T^n} \equiv +1 \mid u_r = -1] \leq \mathbf{P}[u_{T^n} \equiv +1 \mid u_r = +1]$$

Mit der zweiten Ungleichung aus (ii) erhalten wir:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbf{P}[u_{T^n} \equiv +1] &= \frac{1}{2} \cdot \lim_{n \rightarrow \infty} \mathbf{P}[u_{T^n} \equiv +1 \mid u_r = +1] \\ &\quad + \frac{1}{2} \cdot \lim_{n \rightarrow \infty} \mathbf{P}[u_{T^n} \equiv +1 \mid u_r = -1] \\ &\leq \lim_{n \rightarrow \infty} (1 - \epsilon)^{\#T^n} = 0 \end{aligned}$$

(iv) Die Aussage folgt unmittelbar aus (3.11), wo die gegenseitige Information und der D_χ -Abstand miteinander verglichen werden.

Wir setzen $A := \sigma$ und $S : \mathcal{U}^\sigma \rightarrow \{-1, +1\}$ als die Identität. Damit sind ν, ν_-, ν_+ festgelegt. Für $x \in \{-1, +1\}$ ist:

$$\nu_+(x) = \mathbf{P}[u_\sigma = x \mid u_r = +1]$$

$$\nu_-(x) = \mathbf{P}[u_\sigma = x \mid u_r = -1]$$

$$\nu(x) = \mathbf{P}[u_\sigma = x]$$

Mit dieser Wahl ist:

$$D_I(\nu_+, \nu_-) = \mathbf{I}[u_r, u_\sigma]$$

Andererseits zeigt eine elementare Rechnung:

$$D_\chi(\nu_+, \nu_-) = \mathbf{E}[u_r u_\sigma]$$

Mit der Ungleichung (3.11) folgt:

$$\mathbf{I}[u_r, u_\sigma] = D_I(\nu_+, \nu_-) \leq (D_\chi(\nu_+, \nu_-))^2 = \mathbf{E}[u_r u_\sigma]^2 = \theta^{2|\sigma|}$$

□

4.2 Der Interaktionsparameter α

Wie wir bereits gesehen haben, beschreibt die Fehlerwahrscheinlichkeit ϵ die Interaktion in dem Modell. Dabei entsprechen kleine ϵ -Werte starker Interaktion und große ϵ -Werte geringer Interaktion. Wir führen nun einen Interaktionsparameter α ein, der eindeutig mit ϵ zusammenhängt.

Alle Sätze, die eine Aussage über eine Fehlerwahrscheinlichkeit ϵ machen, können demnach auch als eine Aussage über einen Interaktionsparameter α formuliert werden. Wir werden meist beide äquivalenten Aussagen aufführen. Der Grund für diese Ausführlichkeit liegt darin, daß die Aussagen über die Fehlerwahrscheinlichkeit ϵ eine leichte Interpretation im Bild der Signalübertragung finden, die Formulierung als eine Aussage über die Interaktionsstärke α jedoch den Vergleich verschiedener kritischer Parameter vereinfacht.

Dieser Vergleich motiviert, wie wir noch sehen werden, auch folgende

Definition 19. Sei $\frac{1}{2} \geq \epsilon > 0$ die Fehlerwahrscheinlichkeit im Broadcasting-Modell auf T . Dann ist die durch ϵ eindeutig festgelegte **Interaktionsstärke** α definiert als:

$$\alpha(\epsilon) := \frac{1 - \epsilon}{\epsilon} \quad \Longleftrightarrow \quad \epsilon(\alpha) := \frac{1}{1 + \alpha} \quad (4.11)$$

α wird im folgenden auch als **Interaktionsparameter** oder schlicht **Interaktion** bezeichnet.

In Abbildung 4.2 sieht man die grafische Darstellung des Zusammenhangs (4.11). α fällt streng monoton in ϵ und es gilt:

$$\lim_{\epsilon \rightarrow 0} \alpha(\epsilon) = \infty$$

$$\lim_{\epsilon \rightarrow \frac{1}{2}} \alpha(\epsilon) = 1$$

Es gilt $\alpha \in [1, \infty)$, wobei $\alpha = 1$ dem entkoppelten Fall entspricht.

Die lokale Charakteristik läßt sich als Funktion der Interaktionsstärke α ausdrücken. Die Kombination von (4.11) und (4.9) liefert:

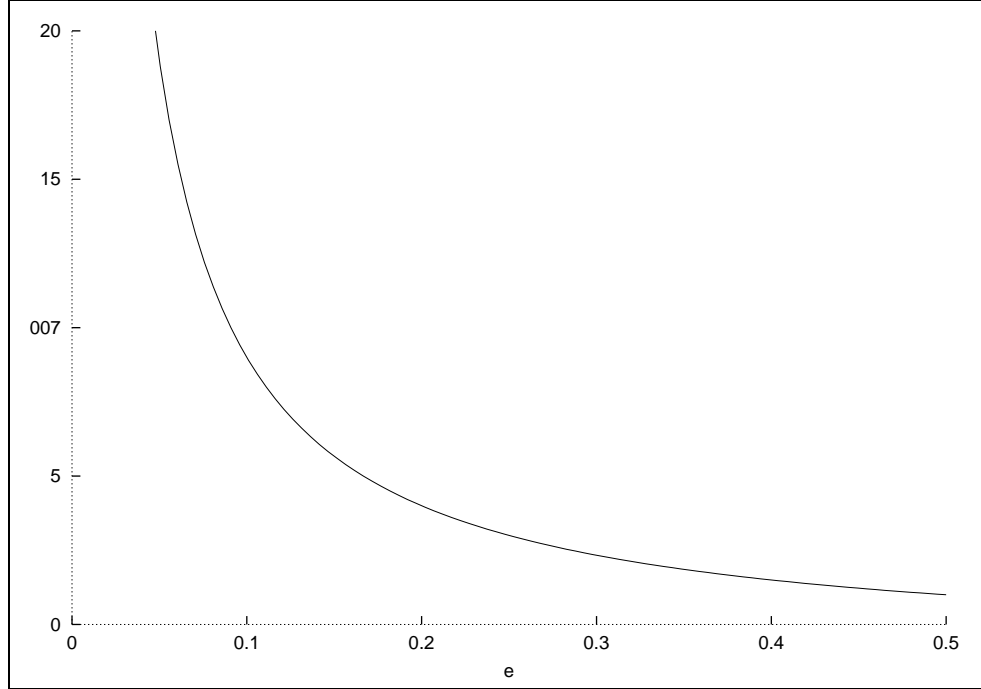


Abbildung 4.2: α als Funktion von ϵ

Korollar 3. *Im Broadcasting-Modell auf T mit Interaktionsstärke α gilt für alle $\sigma \in K_T$:*

$$\begin{aligned} \mathbf{P}[u_\sigma = x \mid u_{K_T \setminus \sigma} = a] &= \frac{\alpha^{l(x,y)-o(x,y)}}{\alpha^{o(x,y)-l(x,y)} + \alpha^{l(x,y)-o(x,y)}} \\ &= \left(1 + \alpha^{o(x,y)-l(x,y)}\right)^{-1} \end{aligned} \quad (4.12)$$

Dabei ist $x \in \{-1, +1\}$, $a \in \{-1, +1\}^{K_T \setminus \sigma}$ und $y = a|_{N(\sigma)}$.

4.3 Kritischer Parameter für Informationsfluß

Wie wir bereits gesehen haben, ist die Fehlerwahrscheinlichkeit ϵ ein Maß dafür, wie stark die Interaktion zwischen benachbarten Knoten ist. In unserem Modell stellt sich die Frage, wie stark die Interaktion mindestens sein muß, damit weit entfernte Gebiete noch Einfluß aufeinander nehmen. Diese

vage Formulierung soll im Rahmen dieser Arbeit auf zwei Arten konkretisiert werden. Wir werden dabei sehen, daß es unterschiedlich scharfe Versionen des Begriffs *Einfluß nehmen* gibt, zu denen auch unterschiedliche Mindestinteraktionen gehören.

Wie schon der Titel dieses Abschnitts vermuten läßt, werden wir uns hier mit der gegenseitigen Information als Maß des Einflusses beschäftigen. Es gilt:

Satz 12. *Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Broadcasting-Modell auf T mit Interaktionsstärke α . T^n sei die n -Sphäre von T . Dann gilt für die Information zwischen Spin der Wurzel und den Spins der n -Sphäre:*

$$\lim_{n \rightarrow \infty} I[u_r, u_{T^n}] = 0, \quad \text{falls } \alpha < \alpha_{c,I} \quad (4.13)$$

$$\lim_{n \rightarrow \infty} I[u_r, u_{T^n}] > 0, \quad \text{falls } \alpha > \alpha_{c,I} \quad (4.14)$$

wobei

$$\alpha_{c,I} := \frac{\sqrt{\text{br}T} + 1}{\sqrt{\text{br}T} - 1} \quad (4.15)$$

ist.

Es gibt im Broadcasting-Modell eine kritische Interaktionsstärke $\alpha_{c,I}$ für Informationsfluß zwischen Wurzel und der unendlich weit entfernten Krone des Baumes. Nach Lemma 2.(ii) ist diese Interaktion auch kritisch für die stochastische Unabhängigkeit von Spin der Wurzel und den Spins der n -Sphäre.

Der Beweis dieses zentralen Satzes befindet sich im nächsten Kapitel. Es handelt sich dabei um den Beweis aus [11], der mit Hilfe einer oberen und einer unteren Abschätzung für $I[u_r, u_{T^n}]$ zeigt, daß $\alpha_{c,I}$ kritisch für Informationsfluß ist.

Durch Einsetzen des Wertes aus Gleichung (4.15) in (4.11) erhalten wir die für Informationsfluß kritische Fehlerwahrscheinlichkeit, wie sie in [11] berechnet wird:

Korollar 4. *Die kritische Fehlerwahrscheinlichkeit für Informationsübertragung zwischen Wurzel und Krone eines Baumes im Sinne von Satz 12 ist.*

$$\epsilon_{c,I} := \frac{\sqrt{\text{br}T} - 1}{2 \cdot \sqrt{\text{br}T}} \quad (4.16)$$

Man bedenke, daß ϵ monoton *fallend* in α ist.

4.4 Das Ising-Modell

Wie bereits angedeutet gibt es eine gänzlich andere Darstellungsmöglichkeit des Broadcasting-Modells. Es ist dies das Ising-Modell auf T , wie es zum Beispiel in [7] betrachtet wird. (Die in [7] auftretenden Konstanten J und k , die für die von uns betrachteten Phänomene keine Bedeutung haben, haben wir zu 1 normiert).

Gegeben sei ein Baum T und eine als **Temperatur** bezeichnete positive reelle Zahl t . Den Knoten des Baumes sind $\{-1, +1\}$ -wertige Zufallsvariablen zugeordnet. Außerdem sei $c : K_T \times K_T \rightarrow \{0, 1\}$ definiert als:

$$\text{mit} \quad c(\sigma, \tau) = \begin{cases} 1 & \text{falls } \tau \text{ und } \sigma \text{ benachbart sind} \\ 0 & \text{sonst} \end{cases}$$

Wir definieren zunächst, was die Energie einer Konfiguration sein soll.

Sei $A \subset \mathcal{C}(T)$ und $a \in \mathcal{U}^A$. Dann ist die Energie von a definiert als:

$$H(a) := \frac{1}{2t} \sum_{\sigma, \tau \in A} a|_{\tau} \cdot a|_{\sigma} \cdot c(\sigma, \tau) \quad (4.17)$$

H ist ein nächstes Nachbarschaftspotential auf T , denn nur benachbarte Spins tragen zur Energie einer Konfiguration bei. H ist demnach geeignet, um ein Gibbs-Maß zu definieren, was wir auch tun werden. Zunächst wollen wir aber noch zwei Eigenschaften von H erwähnen.

Vergleicht man die Energie von zwei Konfigurationen, die sich nur dadurch unterscheiden, daß die Vorzeichen der Spins vertauscht sind, sieht man, daß H ihnen die gleiche Energie gibt. H ist damit ein *Ising-Potential* (vgl. [10] S. 57 oder [3], S. 50). Weiterhin wird H maximal, wenn alle Knoten in A gleichen Spin haben. Das Potential modelliert in diesem Sinne ein *attraktives* System. Dieser Tendenz wirkt der Einfluß der Temperatur entgegen.

Die Baumstruktur ermöglicht es uns, die Energie von Konfigurationen auf $\leq T^n$ kürzer aufzuschreiben. Für $a \in \mathcal{U}^{\leq T^n}$ gilt:

$$H(a) = \frac{1}{t} \sum_{r \neq \sigma \in \leq T^n} a|_{\sigma} \cdot a|_{\tau}$$

Wir werden nun das Ising-Modell mit freien Randbedingungen auf T einführen. Dabei gehen wir vor wie in Lemma 1.

Zunächst definieren wir zu jeder n -Sphäre T^n ein Wahrscheinlichkeitsmaß \mathbf{P}^n und erhalten dann das Gibbs-Maß auf T als Grenzwert dieser Folge von Wahrscheinlichkeitsmaßen.

Definition 20. Gegeben sei ein Baum T und eine als Temperatur bezeichnete reelle Zahl $t > 0$. Jedem Knoten $\sigma \in K_T$ wird eine $\{-1, +1\}$ -wertige Zufallsvariable u_σ zugeordnet. \mathcal{A} sei die von den endlichdimensionalen Zylindermengen erzeugte σ -Algebra auf \mathcal{U}^{K_T} . Mit H aus (4.17) definieren wir zu jedem T^n ein Wahrscheinlichkeitsmaß \mathbf{P}^n wie folgt:

$$\mathbf{P}^n[u_{\leq T^n} = a] := \frac{\exp[H(a)]}{\sum_{c \in \mathcal{U}^{\leq T^n}} \exp[H(c)]} , \quad \text{für } a \in \mathcal{U}^{\leq T^n}$$

$$\mathbf{P}^n[u_\sigma = -1] := 1 , \quad \text{für } |\sigma| > n$$

Nach Lemma 1 ist $\mathbf{P}' := \lim_{n \rightarrow \infty} \mathbf{P}^n$ ein Gibbs-Maß auf T mit Potential H . Wir nennen $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P}')$ **Ising-Modell auf T mit Temperatur t** .

Wir können dem Ising-Modell eine anschauliche Interpretation in einem physikalischen Rahmen geben. Die Zufallsvariablen der Knoten repräsentieren magnetische Spins, die einen von zwei Zuständen annehmen können. Nur direkt benachbarte Knoten beeinflussen sich in dem Sinn, daß ein Spin umso eher einen Zustand annimmt, je mehr seiner Nachbarn diesen Zustand haben. Diesem Einfluß wirkt die Temperatur entgegen.

Wie schon in Gleichung (4.11) ist es auch hier möglich, die Temperatur des Ising-Modells eindeutig in eine Interaktionsstärke α' umzurechnen. Dies wird uns bald hilfreich sein, um Broadcasting-Modell und Ising-Modell miteinander zu vergleichen.

Definition 21. Im Ising-Modell auf T definieren wir die durch die Temperatur eindeutig festgelegte Interaktionsstärke α' als:

$$\alpha'(t) := \exp[2/t] \quad \Longleftrightarrow \quad t(\alpha') := \frac{2}{\ln(\alpha')} \quad (4.18)$$

Für die im Ising-Modell zulässigen Temperaturen ist $\alpha' \in [1, \infty)$. Der funktionale Zusammenhang zwischen α' und t ist in Bild 4.3 dargestellt. α' fällt

streng monoton als Funktion von t , und:

$$\lim_{t \rightarrow 0} \alpha'(t) = \infty \quad (4.19)$$

$$\lim_{t \rightarrow \infty} \alpha'(t) = 1 \quad (4.20)$$

Hohe Temperaturen entsprechen geringer Interaktion zwischen den Spins. Dies paßt zu unserer physikalischen Interpretation. Bei hohen Temperaturen überlagert die thermische Unordnung den gegenseitigen Einfluß der Spins.

Das Wahrscheinlichkeitsmaß \mathbf{P}' ist auf Grund seiner Konstruktion ein Gibbs-Maß. Wegen des elementaren Zusammenhangs aus Satz 6 ist klar, daß \mathbf{P}' auch ein Markov-Feld auf \mathcal{U}^{K_T} ist. Im folgenden Satz wird dies konkretisiert.

Satz 13. *Das in Definition 20 beschriebene Wahrscheinlichkeitsmaß \mathbf{P}' ist ein Markov-Feld auf K_T mit der lokalen Charakteristik:*

$$\mathbf{P}'[u_\sigma = x \mid u_{K_T \setminus \sigma} = a] = \mathbf{P}'[u_\sigma = x \mid u_{N(\sigma)} = y] \quad (4.21)$$

$$\begin{aligned} &= \frac{\exp[1/t]^{l(x,y)-o(x,y)}}{\exp[1/t]^{l(x,y)-o(x,y)} + \exp[1/t]^{o(x,y)-l(x,y)}} \\ &= \left(1 + \exp[2/t]^{o(x,y)-l(x,y)}\right)^{-1} \end{aligned} \quad (4.22)$$

Dabei ist $x \in \{-1, +1\}$, $a \in \{-1, +1\}^{K_T \setminus \sigma}$ und $y = a|_{N(\sigma)}$. $l(x, y)$ und $o(x, y)$ sind definiert wie in (4.3) und (4.4).

Beweis. Wegen des elementaren Zusammenhangs zwischen Gibbs-Maßen und Markov-Feldern ist klar, daß Gleichung (4.21) gilt. Es bleibt zu zeigen, daß das Ising-Modell die angegebene lokale Charakteristik hat.

Sei $b = (x, a) \in \mathcal{U}^{K_T}$ die Belegung mit $b|_\sigma = x$ und $b|_{K_T \setminus \sigma} = a$, analog $\bar{b} = (-x, a) \in \mathcal{U}^{K_T}$. Insbesondere ist also $b|_{(K_T \setminus \sigma)} = \bar{b}|_{(K_T \setminus \sigma)}$. Mit diesen Be-

zeichnungen gilt:

$$\begin{aligned}
\mathbf{P}'[u_\sigma = x \mid u_{K_T \setminus \sigma} = a] &= \\
&= \frac{\mathbf{P}'[u_\sigma = x, u_{K_T \setminus \sigma} = a]}{\mathbf{P}'[u_{K_T \setminus \sigma} = a]} \\
&= \frac{\mathbf{P}'[u_\sigma = x, u_{K_T \setminus \sigma} = a]}{\mathbf{P}'[u_\sigma = x, u_{K_T \setminus \sigma} = a] + \mathbf{P}'[u_\sigma = -x, u_{K_T \setminus \sigma} = a]} \\
&= \frac{\exp[-H[u_T = b]]}{\exp[-H[u_T = b]] + \exp[-H[u_T = \bar{b}]]} \\
&= \frac{\exp\left[1/t \sum_{\sigma \in T} b_\sigma b_{\bar{\sigma}}\right]}{\exp\left[1/t \sum_{\sigma \in T} b_\sigma b_{\bar{\sigma}}\right] + \exp\left[1/t \sum_{\sigma \in T} \bar{b}_\sigma \bar{b}_{\bar{\sigma}}\right]} \tag{4.23}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\exp\left[1/t \left(\sum_{\tau \in N(\sigma)} b_\sigma b_\tau\right)\right]}{\exp\left[1/t \left(\sum_{\tau \in N(\sigma)} b_\sigma b_\tau\right)\right] + \exp\left[1/t \left(\sum_{\tau \in N(\sigma)} \bar{b}_\sigma \bar{b}_\tau\right)\right]} \tag{4.24} \\
&= \frac{\exp[1/t]^{l(x,y)-o(x,y)}}{\exp[1/t]^{l(x,y)-o(x,y)} + \exp[1/t]^{o(x,y)-l(x,y)}} \\
&= \left(1 + \exp[2/t]^{o(x,y)-l(x,y)}\right)^{-1}
\end{aligned}$$

Die Umformung von (4.23) nach (4.24) benutzt die Funktionalgleichung der Exponentialfunktion und die Eigenschaft $b|_{(K_T \setminus \sigma)} = \bar{b}|_{(K_T \setminus \sigma)}$. \square

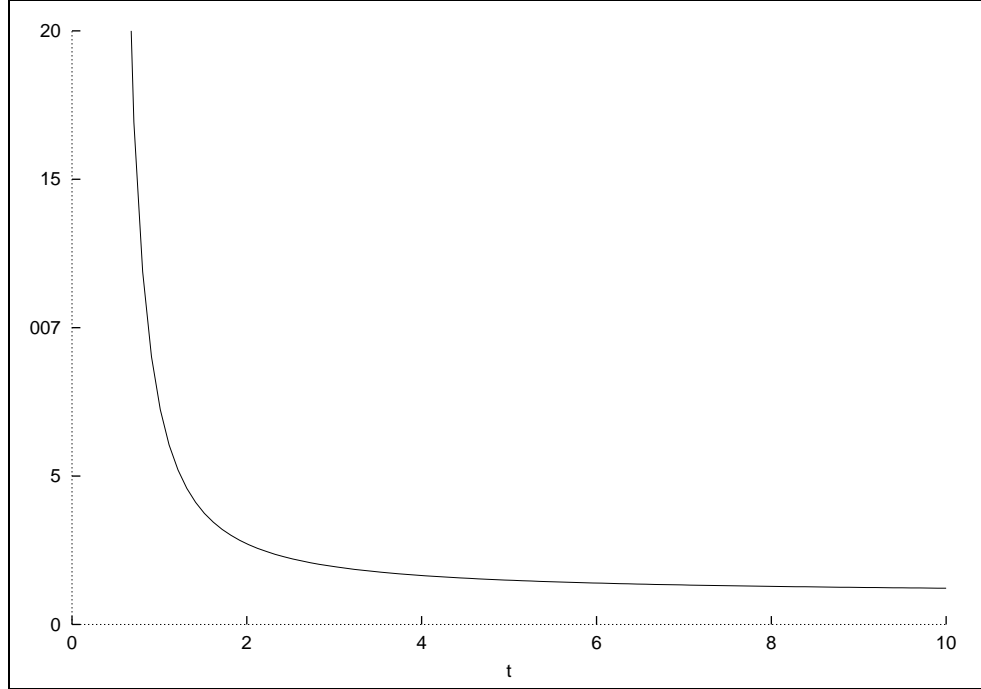


Abbildung 4.3: Interaktionsstärke α' als Funktion der Temperatur t

Natürlich läßt sich auch im Ising-Modell die lokale Charakteristik durch die Interaktionsstärke α' ausdrücken. Einsetzen des Zusammenhangs (4.18) in Gleichung (4.22) liefert:

Korollar 5. *Im Ising-Modell mit Interaktionsstärke α' gilt mit den Bezeichnungen aus Satz 13 für alle $\sigma \in K_T$:*

$$\begin{aligned} \mathbf{P}'[u_\sigma = x \mid u_{K_T \setminus \sigma} = a] &= \frac{\alpha'^{l(x,y)-o(x,y)}}{\alpha'^{o(x,y)-l(x,y)} + \alpha'^{l(x,y)-o(x,y)}} \\ &= \left(1 + \alpha'^{2(o(x,y)-l(x,y))}\right)^{-1} \end{aligned} \quad (4.25)$$

Der Vergleich der lokalen Charakteristika in Korollar 3 und Korollar 5 legt die Vermutung nahe, daß das Ising-Modell mit freien Randbedingungen und das Broadcasting-Modell identisch sind. Dies trifft tatsächlich zu:

Satz 14. *Broadcasting-Modell und Ising-Modell sind identisch, wenn die Modellparameter Temperatur t und Fehlerwahrscheinlichkeit ϵ wie folgt umgerechnet werden.*

$$\epsilon(t) = \frac{1}{1 + \exp\left[\frac{2}{t}\right]} \quad \Longleftrightarrow \quad t(\epsilon) = \frac{2}{\ln\left[\frac{1-\epsilon}{\epsilon}\right]} \quad (4.26)$$

Beweis. Wir zeigen, daß die Überführung der Parameter gemäß (4.26) bewirkt, daß die Wahrscheinlichkeiten endlicher Konfigurationen unter \mathbf{P} und \mathbf{P}' proportional zueinander sind.

Wir betrachten zunächst Konfigurationen auf $\leq T^n$. Sei dazu $n \in \mathbb{N}$ und $a \in \mathcal{U}^{\leq T^n}$ fest gewählt. N sei die Anzahl der Kanten des Baumes, die zwischen Wurzel und der n -Sphäre liegen. Weiter seien

$$l := \frac{1}{2} \cdot \#\{\tau, \sigma \in \leq T^n : a|_\sigma = a|_\tau\}$$

$$o := \frac{1}{2} \cdot \#\{\tau, \sigma \in \leq T^n : a|_\sigma \neq a|_\tau\}$$

Damit ist $N = l + o$. Mit diesen Bezeichnungen gilt:

$$\begin{aligned} \mathbf{P}[u_{\leq T^n} = a] &= \frac{1}{2} \cdot (1 - \epsilon)^l \epsilon^o \\ &= \frac{1}{2} \frac{\left(\frac{1-\epsilon}{\epsilon}\right)^l}{\left(\frac{1}{\epsilon}\right)^N} \end{aligned}$$

Die Energie von a im Ising-Modell ist:

$$H(a) = \frac{1}{t}(l - o)$$

Mit dem Wert von ϵ aus (4.26) folgt:

$$\begin{aligned} \exp[H(a)] &= \exp[1/t]^l \cdot \exp[-1/t]^o \\ &= \frac{\left(\frac{1-\epsilon}{\epsilon}\right)^l}{\left(\frac{1-\epsilon}{\epsilon}\right)^{N/2}} \end{aligned}$$

Wenn wir noch die Normierungskonstante Z_n als

$$Z_n := \sum_{b \in \mathcal{U}^{\leq T^n}} \exp[H(b)]$$

einführen, können wir die Wahrscheinlichkeit der Konfiguration a im Ising-Modell wie folgt darstellen:

$$\begin{aligned}\mathbf{P}'[u_{\leq T^n} = a] &= \frac{\exp[H(a)]}{Z_n} \\ &= \frac{\left(\frac{1-\epsilon}{\epsilon}\right)^l}{\left(\frac{1-\epsilon}{\epsilon}\right)^{N/2} \cdot Z_n}\end{aligned}$$

Für das Verhältnis der Wahrscheinlichkeiten der Konfiguration a in den beiden Modellen gilt nun:

$$\begin{aligned}\frac{\mathbf{P}[u_{\leq T^n} = a]}{\mathbf{P}'[u_{\leq T^n} = a]} &= \frac{1}{2} \frac{\left(\frac{1-\epsilon}{\epsilon}\right)^l}{\left(\frac{1}{\epsilon}\right)^N} \cdot \frac{\left(\frac{1-\epsilon}{\epsilon}\right)^{N/2} \cdot Z_n}{\left(\frac{1-\epsilon}{\epsilon}\right)^l} \\ &= \frac{1}{2} ((1-\epsilon)\epsilon)^{N/2} \cdot Z_n =: c_n\end{aligned}\tag{4.27}$$

c_n ist nicht von der Konfiguration abhängig, Konfigurationen auf $\leq T^n$ haben also unter \mathbf{P} und \mathbf{P}' proportionale Wahrscheinlichkeiten. Da \mathbf{P} und \mathbf{P}' Wahrscheinlichkeitsmaße sind folgt aus 4.27:

$$\begin{aligned}1 &= \sum_{a \in \mathcal{U}^{\leq T^n}} \mathbf{P}[u_{\leq T^n} = a] \\ &= c_n \cdot \sum_{a \in \mathcal{U}^{\leq T^n}} \mathbf{P}'[u_{\leq T^n} = a] \\ &= c_n \cdot 1\end{aligned}$$

Dies bedeutet, daß $c_n = 1$ die Wahrscheinlichkeiten von Konfigurationen auf $\leq T^n$ bezüglich \mathbf{P} und \mathbf{P}' sogar gleich sind, wenn wir die Parameter gemäß (4.26) ineinander überführen.

Sei nun $A \in \mathcal{C}(T)$ eine beliebige endliche Menge von Knoten und $a \in \mathcal{U}^A$ eine Konfiguration auf A . Weiterhin sei $m \in \mathbb{N}$ so gewählt, daß $A \subset T^m$ ist. Dann

gilt:

$$\begin{aligned}
\mathbf{P}[u_A = a] &= \sum_{b \in \mathcal{U}^{\leq T^m}, b|_A = a} \mathbf{P}[u_{\leq T^m} = b] \\
&= \sum_{b \in \mathcal{U}^{\leq T^m}, b|_A = a} \mathbf{P}'[u_{\leq T^m} = b] \\
&= \mathbf{P}'[u_A = a]
\end{aligned}$$

Die Wahrscheinlichkeiten aller endlichen Konfigurationen sind demnach unter \mathbf{P} und \mathbf{P}' gleich, wenn wir die Parameter gemäß (4.26) ineinander überführen. Nach Korollar 1 werden \mathbf{P} und \mathbf{P}' durch die Gewichte solcher Konfigurationen festgelegt, und sind damit gleich. \square

Es gibt also keine zwei Modelle, sondern lediglich zwei *Darstellungen* eines Modells. Das Broadcasting-Modell liefert eine neue Darstellung des Ising-Modells mit freien Randbedingungen. Wir werden in einigen folgenden Rechnungen sehen, daß die Beschreibung als Broadcasting-Modell einige Vorteile bietet, die in der Unabhängigkeit der Zufallsvariablen auf den Kanten begründet ist. Besonders bei der Untersuchung von Spinperkolation wird uns dies sehr hilfreich sein.

Von nun ab unterscheiden wir nicht mehr zwischen \mathbf{P} und \mathbf{P}' sowie zwischen α und α' , und wechseln die Darstellung je nach betrachtetem Aspekt.

4.5 Kritischer Parameter für Phasenübergang

Eine wesentliche Frage im Ising-Modell ist, ob Phasenübergang vorliegt. Ein sicheres Indiz für Phasenübergang sind nach Satz 8 unterschiedliche Wahrscheinlichkeiten des Spins in der Wurzel in der Plus- und in der Minus-Phase. Das können wir als Einfluß der $+1$ bzw. -1 -Randbedingungen am Rand von T interpretieren. Dies ist die zweite Version des Begriffs *Einfluß haben*.

Wir definieren Minus- und Plus-Phase des Ising-Modells. Dabei gehen wir vor wie bereits in Definition 20.

Definition 22. Seien T, \mathcal{A} und H definiert wie in Definition 20. Für jedes T^n definieren wir:

$$\mathbf{P}^{n,+}(u_{<T^n} = a) := \frac{\exp[H(a, u_{T^n} \equiv +1)]}{\sum_{c \in \mathcal{U}^{<T^n}} \exp[H(c, u_{T^n} \equiv +1)]}, \quad \text{für } a \in \mathcal{U}^{<T^n}$$

$$\mathbf{P}^{n,+}[u_\sigma = +1] := 1, \quad \text{für } \sigma \in T^n$$

$$\mathbf{P}^{n,+}[u_\sigma = -1] := 1, \quad \text{für } \sigma \notin \leq T^n$$

$$\mathbf{P}^{n,-}(u_{<T^n} = a) := \frac{\exp[H(a, u_{T^n} \equiv -1)]}{\sum_{c \in \mathcal{U}^{<T^n}} \exp[H(c, u_{T^n} \equiv +1)]}, \quad \text{für } a \in \mathcal{U}^{<T^n}$$

$$\mathbf{P}^{n,-}[u_\sigma = -1] := 1, \quad \text{für } \sigma \in T^n$$

$$\mathbf{P}^{n,-}[u_\sigma = +1] := 1, \quad \text{für } \sigma \notin \leq T^n$$

Nach Lemma 1 gilt: $\lim_{n \rightarrow \infty} \mathbf{P}^{n,-} =: \mathbf{P}^- \in \mathcal{G}_H$ und $\lim_{n \rightarrow \infty} \mathbf{P}^{n,+} =: \mathbf{P}^+ \in \mathcal{G}_H$. Wir nennen $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P}^-)$ **Minus-Phase** des Ising-Modells und $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P}^+)$ **Plus-Phase** des Ising-Modells.

Wir können die oben auftauchenden Wahrscheinlichkeitsmaße $\mathbf{P}^{n,+}$ und $\mathbf{P}^{n,-}$ auch anders darstellen. Mit \mathbf{P}^n aus Definition 20 gilt für $a \in \mathcal{U}^{<T^n}$:

$$\mathbf{P}^{n,+}[u_{<T^n} = a] = \mathbf{P}^n[u_{<T^n} = a \mid u_{T^n} \equiv +1] \quad (4.28)$$

$$\mathbf{P}^{n,-}[u_{<T^n} = a] = \mathbf{P}^n[u_{<T^n} = a \mid u_{T^n} \equiv -1] \quad (4.29)$$

Dies motiviert die Anschauung, daß wir die Plus-Phase des Ising-Modells als das freie Ising-Modell, bedingt darauf, daß die Spins der Krone alle +1 sind, verstehen wollen. Ebenso natürlich die Minus-Phase als freies Ising-Modell bedingt auf die -1-Randbedingung.

Die Darstellung aus (4.28) und (4.29) zeigt uns, daß wir Minus- und Plus-Phase richtig definiert haben, denn aus der Attraktivität des Ising-Modells folgt für alle $y \in \mathcal{U}^{T^n}$ und für alle $A \subset < T^n$:

$$\mathbf{P}^n[u_A \equiv +1 \mid u_{T^n} \equiv -1] \leq \mathbf{P}^n[u_A \equiv +1 \mid u_{T^n} = y] \leq \mathbf{P}^n[u_A \equiv +1 \mid u_{T^n} \equiv +1]$$

Plus- und Minus-Phase sind damit die in \mathcal{G}_H ausgezeichneten Maße, die in Satz 7 erwähnt werden.

Es stellt sich nun die Frage, ob die Wahrscheinlichkeit für den Spin $+1$ in der Wurzel dadurch verändert wird, daß wir darauf bedingen, daß die Spins am Rand $\equiv +1$ sind. Wir fragen also, wann $\mathbf{P}^+[u_r = +1] = \mathbf{P}[u_r = +1]$ ist. Die Antwort gibt der folgende

Satz 15. *Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Ising-Modell mit Interaktionsstärke α . Dann gilt:*

$$\mathbf{P}[u_r = 1] < \mathbf{P}^+[u_r = 1] , \quad \text{falls } \alpha > \alpha_{c,P} \quad (4.30)$$

$$\mathbf{P}[u_r = 1] = \mathbf{P}^+[u_r = 1] , \quad \text{falls } \alpha < \alpha_{c,P} \quad (4.31)$$

Dabei ist

$$\alpha_{c,P} := \frac{\text{brT} + 1}{\text{brT} - 1} \quad (4.32)$$

Wir beweisen diese Aussage im nächsten Kapitel. Dabei werden wir den Beweis so führen, wie Lyons in [7] die kritische Temperatur für Phasenübergang im Ising-Modell berechnet hat.

Wir geben hier noch Lyons Aussage als Korollar an:

Korollar 6. *Im Ising-Modell mit Temperatur t gilt:*

$$\mathbf{P}[u_r = 1] < \mathbf{P}^+[u_r = 1] , \quad \text{falls } t < t_c \quad (4.33)$$

$$\mathbf{P}[u_r = 1] = \mathbf{P}^+[u_r = 1] , \quad \text{falls } t > t_c \quad (4.34)$$

Dabei ist

$$t_c = \frac{1}{\text{arccoth}(\text{brT})} \quad (4.35)$$

Beweis. Der Beweis besteht wieder nur aus dem Einsetzen des Wertes für t_c in (4.18), wobei wir verwenden, daß

$$\text{arccoth}(x) = \frac{1}{2} \log \left[\frac{x+1}{x-1} \right] \quad (4.36)$$

ist. Damit gilt:

$$\begin{aligned}
t_c &= t(\alpha_{c,P}) \\
&= \frac{2}{\log\left(\frac{\text{brT}+1}{\text{brT}-1}\right)} \\
&= \left(\frac{1}{2} \frac{\text{brT}+1}{\text{brT}-1}\right)^{-1} \\
&= \frac{1}{\text{arccoth}(\text{brT})}
\end{aligned}$$

□

Wir interpretieren Satz 15 so, daß für Interaktionsstärken α unterhalb von $\alpha_{c,P}$ die Wahrscheinlichkeit eines Spins in der Wurzel in der Plus-Phase gleich der Wahrscheinlichkeit ohne Vorgabe einer Randbedingung ist. In diesem Sinne hat der Rand für $\alpha < \alpha_{c,P}$ keinen Einfluß auf die Wurzel. Für $\alpha > \alpha_{c,P}$ gibt die Plus-Phase den Spins in der Wurzel eine andere Wahrscheinlichkeit als \mathbf{P} . Wir verstehen dies so, daß die Randbedingung “Rand $\equiv +1$ ” hier Einfluß auf den Spin der Wurzel ausübt.

Laut Satz 8 sind die Aussagen aus Satz 15 oder Korollar 6 eine Aussage über eine kritische Schwelle für Phasenübergang. Dies halten wir in folgendem Korollar fest:

Korollar 7. *Die oben angegebenen kritischen Parameter t_c und $\alpha_{c,P}$ sind kritisch für Phasenübergang im Ising-Modell auf T .*

4.6 Vergleich der kritischen Interaktionsstärken

Nun werden wir die beiden Arten der Einflußnahme des unendlich weit entfernten Randes auf die Wurzel miteinander vergleichen. Dazu stellen wir zunächst die kritischen Interaktionsparameter gegenüber. Dies wird uns dabei helfen, die Unterschiede der Einflußnahmen zu erkennen.

Satz 16. *Im Broadcasting-Modell ist die kritische Interaktionsstärke für Phasenübergang stets echt kleiner als die kritische Interaktionsstärke für Informationsfluß. Es gilt also stets:*

$$\alpha_{c,P} < \alpha_{c,I} \quad (4.37)$$

Beweis.

$$\begin{aligned} \alpha_{c,P} < \alpha_{c,I} &\iff \frac{\text{brT} + 1}{\text{brT} - 1} < \frac{\sqrt{\text{brT}} + 1}{\sqrt{\text{brT}} - 1} \cdot \frac{\sqrt{\text{brT}} + 1}{\sqrt{\text{brT}} + 1} \\ &\iff \frac{\text{brT} + 1}{\text{brT} - 1} < \frac{(\sqrt{\text{brT}} + 1)^2}{\text{brT} - 1} \\ &\iff \text{brT} + 1 < \text{brT} + 1 + 2\sqrt{\text{brT}} \end{aligned}$$

□

Wir wollen die Aussage dieses Satzes ein wenig genauer betrachten und interpretieren. Sei dazu im folgenden ein Baum T fest gewählt. $\alpha_{c,P}$ und $\alpha_{c,I}$ seien die kritischen Interaktionsstärken aus Satz 15 und Satz 12. Wir unterscheiden drei charakteristische Fälle.

Sei zunächst $\alpha < \alpha_{c,P}$. Nach Satz 15 findet für diese Interaktionsstärke kein Phasenübergang statt, Plus- und Minus-Phase sind identisch und damit gilt insbesondere:

$$\mathbf{P}^+[u_r = +1] = \mathbf{P}^-[u_r = +1] = \frac{1}{2}$$

Wir haben in Definition 22 \mathbf{P}^+ und \mathbf{P}^- als die Maße eingeführt, die als Grenzfälle des thermodynamischen Limes für die Randbedingungen $u_{T^n} \equiv +1$ bzw. $u_{T^n} \equiv -1$ entstehen. Wir wollen nun Plus- und Minus-Phase in diesem Sinne als die auf Rand $\equiv +1$ bzw. Rand $\equiv -1$ bedingten Wahrscheinlichkeitsmaße verstehen.

Nach Satz 7 sind \mathbf{P}^+ und \mathbf{P}^- die Wahrscheinlichkeitsmaße, die dem Ereignis $u_r = +1$ die größte bzw. kleinste Wahrscheinlichkeit zuweisen. Im Fall $\alpha < \alpha_{c,P}$ spürt der Spin der Wurzel den Einfluß der extremen Randbedingungen von Plus- bzw. Minus-Phase nicht, und damit auch nicht den Einfluß anderer Randbedingungen.

Der Spin der Wurzel ist von den Spins der Krone des Baumes entkoppelt und dies bedeutet insbesondere, daß keine Information übertragen wird.

Der Fall $\alpha > \alpha_{c,I}$ ist ebenso einsichtig. Damit die Information zwischen Spin der Wurzel und den Spins der Krone positiv ist, dürfen diese nicht unabhängig sein. Es muß demnach eine Belegung auf der Krone geben, so daß sich, gegeben diese Konfiguration, die bedingte Wahrscheinlichkeit für $u_r = +1$ von der bedingten Wahrscheinlichkeit für $u_r = -1$ unterscheidet. Satz 7 besagt nun, daß sich auch die extremen Randbedingungen der Plus-Phase und der Minus-Phase bemerkbar machen. Also ist:

$$\mathbf{P}^+[u_r = +1] \neq \mathbf{P}^-[u_r = -1]$$

Dies ist nach Satz 15 gleichbedeutend zu $\alpha > \alpha_{c,P}$.

Die eigentliche Erkenntnis aus Satz 16 liegt darin, daß $\alpha_{c,P}$ *echt kleiner* als $\alpha_{c,I}$ ist. Es gibt Broadcasting-Modelle, in denen Phasenübergang stattfindet, aber keine Information zwischen Wurzel und Krone übertragen wird. Wie können wir dies verstehen?

Betrachten wir dazu ein Broadcasting-Modell mit Interaktionsstärke α , wobei $\alpha_{c,P} < \alpha < \alpha_{c,I}$ ist. Nach Satz 15 findet für diese Interaktion Phasenübergang statt. Die bedingte Wahrscheinlichkeit für $u_r = +1$ gegeben “Rand $\equiv +1$ ” ist ungleich der bedingten Wahrscheinlichkeit für $u_r = +1$ gegeben “Rand $\equiv -1$ ”.

Andererseits ist bei dieser Interaktionsstärke nach Satz 12 die Information zwischen Spin der Wurzel und der Spinkonfiguration des Randes 0. Der Spin der Wurzel und die Spins der Krone sind unabhängig.

Wir verstehen dies so, daß die extremsten Randbedingungen Rand $\equiv +1$ bzw. Rand $\equiv -1$ zwar durchaus die bedingte Verteilung der Spins der Wurzel beeinflussen, aber diese extremsten Randbedingungen im Broadcasting-Modell nur mit Wahrscheinlichkeit 0 auftreten (wie dies Satz 11 (iii) besagt). Daher tragen sie nichts zur Information zwischen Spin der Wurzel und den Spins des Randes von T bei.

Damit die Information zwischen Spin der Wurzel und den Spins der Krone positiv ist, muß eine Belegung der Krone, die im Broadcasting-Modell mit positiver Wahrscheinlichkeit auftritt, die bedingte Verteilung von u_r beeinflussen.

4.7 Spinperkolation

Wir werden nun das Phänomen der *Spinperkolation* im Broadcasting-Modell untersuchen. Darunter verstehen wir die Existenz eines Pfades, der in der Wurzel startet, bis zum Rand reicht und nur Knoten gleichen Spins besucht. Es stellt sich die Frage, ob die Wahrscheinlichkeit dieses Ereignisses durch die Vorgabe von Randbedingungen beeinflusst wird.

Es ist auf einfache Weise möglich, einen Zusammenhang zwischen einer Spinperkolation im Broadcasting-Modell und einer Perkolation des Baumes T herzustellen. Dieser wird dadurch geschaffen, daß man in T alle die Kanten entfernt, die zwei Knoten unterschiedlichen Spins verbinden.

Definition 23. Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Broadcasting-Modell auf T . Für jede Belegung $a \in \mathcal{U}^{K_T}$ sei:

$$\hat{E}(a) := \{e_{(\sigma, \bar{\sigma})} \in E_T : a|_{\sigma} = a|_{\bar{\sigma}}\} \quad \text{und} \quad \hat{T}(a) := (K_T, \hat{E}(a))$$

Weiterhin sei

$$\mathcal{T} := \{\hat{T}(a) : a \in \mathcal{U}^{K_T}\}$$

Die so erhaltene Perkolation $(\mathcal{T}, \mathbf{P})$ heißt **freie Spinperkolation des Broadcasting-Modells** auf T . Entsprechend sind $(\mathcal{T}, \mathbf{P}^+)$ und $(\mathcal{T}, \mathbf{P}^-)$ mit \mathbf{P}^+ und \mathbf{P}^- aus 22 als **Spinperkolation in der Plus- bzw. Minus-Phase** des Broadcasting-Modells definiert.

In Bild 4.4 wird eine Realisierung einer Spinperkolation gezeigt. Die Frage, ob es im Broadcasting-Modell einen Pfad gibt, der in der Wurzel beginnt, unendlich lang ist und nur Knoten gleichen Spins besucht, wird mit Definition 23 nun zu der Frage, ob in \hat{T} die Zusammenhangskomponente der Wurzel, die wir als \hat{T}_r bezeichnen, unendlich groß ist. Dies ist die Frage danach, ob in \hat{T} *Perkolation stattfindet*. Wir werden im besonderen untersuchen, ob die Vorgabe von Randbedingungen die Wahrscheinlichkeit eines unendlichen Pfades in \hat{T}_r verändert.

Wenn wir im Broadcasting-Modell keine Randbedingungen vorgeben, ist die Spinperkolation eine Bernoulli-Perkolation, denn es werden genau die Kanten

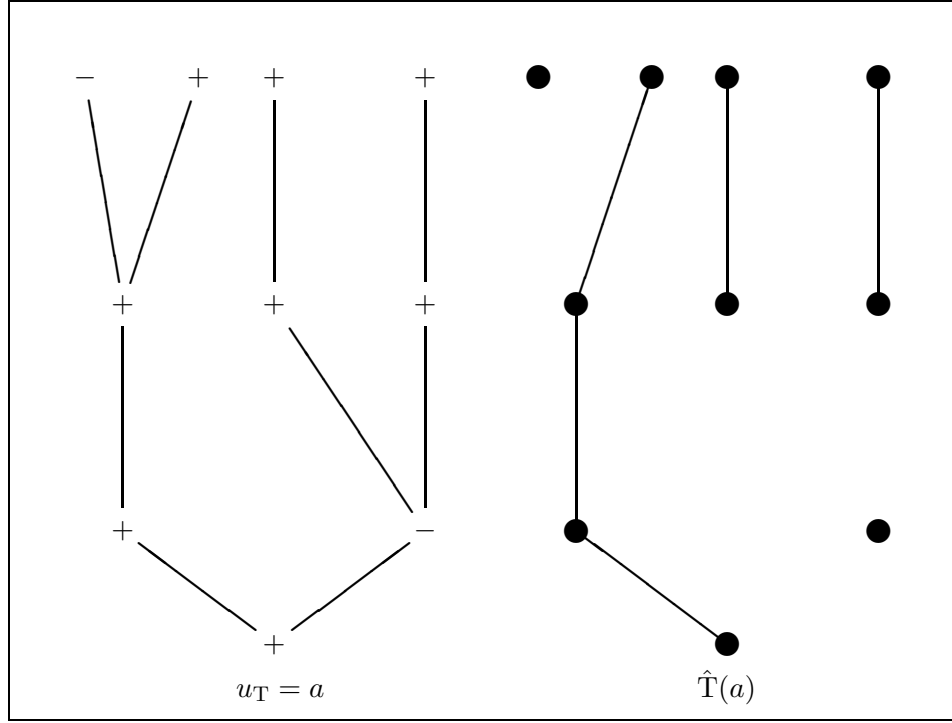


Abbildung 4.4: Eine Belegung auf T und die zugehörige Spinperkolation

entfernt, die fehlerhaft übertragen. Dies geschieht unabhängig an allen Kanten mit der Wahrscheinlichkeit ϵ . Dank Satz 5 können wir sofort eine kritische Fehlerwahrscheinlichkeit $\epsilon_{c,Perk.}$ für Spinperkolation im Broadcasting-Modell ohne Randbedingungen angeben:

Satz 17. *Sei (T, \mathbf{P}) die freie Spinperkolation im Broadcasting-Modell mit Fehlerwahrscheinlichkeit ϵ . Wie in Definition 7 bezeichnen wir die Zusammenhangskomponente der Wurzel in \hat{T} als \hat{T}_r . Dann gilt*

$$\mathbf{P}[|\hat{T}_r| = \infty] = 0, \quad \text{für } \epsilon > \epsilon_{c,Perk.}$$

$$\mathbf{P}[|\hat{T}_r| = \infty] = 1, \quad \text{für } \epsilon < \epsilon_{c,Perk.}$$

Dabei ist

$$\epsilon_{c,Perk.} = \begin{cases} \frac{\text{br}T-1}{\text{br}T} & \text{falls } \text{br}T < 2 \\ \frac{1}{2} & \text{falls } \text{br}T \geq 2 \end{cases}$$

Beweis. Die freie Spinperkolation im Broadcasting-Modell ist eine Bernoulli-Perkolations. Dabei ist die Überlebenswahrscheinlichkeit jeder Kante:

$$p_s = \mathbf{P}[e \in \hat{E}(u)] = \mathbf{P}[\eta_e = 1] = 1 - \epsilon$$

Gemäß Satz 5, der eine Aussage über die kritische Überlebenswahrscheinlichkeit für Bernoulli-Perkolationen macht, folgt:

$$\epsilon_{c,Perk.} = 1 - p_c = 1 - \frac{1}{\text{brT}} = \frac{\text{brT} - 1}{\text{brT}} \quad (4.38)$$

$\epsilon_{c,Perk.}$ ist kritisch für Perkolations in \hat{T} und somit auch kritisch für Spinperkolations im Broadcasting-Modell auf T .

Da wir im Broadcasting-Modell ausschließlich Fehlerwahrscheinlichkeiten $\epsilon < \frac{1}{2}$ betrachten, folgt aus Gleichung (4.38) unmittelbar die fast-sichere Existenz von Spinperkolations für Bäume, deren Branching-Number 2 oder größer ist. \square

$\epsilon_{c,Perk.}$ läßt sich wieder in einen kritischen Interaktionsparameter $\alpha_{c,Perk.}$ umrechnen. Es gilt:

Korollar 8. *Sei (T, \mathbf{P}) die freie Spinperkolations im Broadcasting-Modell mit Interaktionsstärke α . Dann gilt:*

$$\begin{aligned} \mathbf{P}[|\hat{T}_r| = \infty] &= 0, & \text{falls } \alpha < \alpha_{c,Perk.} \\ \mathbf{P}[|\hat{T}_r| = \infty] &= 1, & \text{falls } \alpha > \alpha_{c,Perk.} \end{aligned}$$

Dabei ist

$$\alpha_{c,Perk.} = \begin{cases} \frac{1}{\text{brT}-1} & \text{falls } \text{brT} < 2 \\ 1 & \text{falls } \text{brT} \geq 2 \end{cases}$$

Bevor wir untersuchen, ob die Randbedingungen der Plus-Phase Einfluß auf die Existenz von unendlichen Pfaden gleichen Spins ausüben, wollen wir die kritischen Interaktionen, die wir bisher berechnet haben, miteinander vergleichen.

Satz 18. *Im Broadcasting-Modell auf T gilt für die kritischen Interaktionsparameter:*

$$\alpha_{c,Perk.} < \alpha_{c,P} < \alpha_{c,I} \quad (4.39)$$

Beweis. Nach Definition der kritischen Parameter ist Gleichung (4.39) für $\text{brT} < 2$ äquivalent zu:

$$\frac{1}{\text{brT} - 1} < \frac{\text{brT} + 1}{\text{brT} - 1} < \frac{\sqrt{\text{brT}} + 1}{\sqrt{\text{brT}} - 1}$$

Die erste Ungleichung gilt, da $\text{brT} > 1$ ist, die zweite Ungleichung haben wir bereits in Satz 16 bewiesen.

Für $\text{brT} \geq 2$ gilt 4.39 ebenfalls, denn hier ist $\alpha_{c, \text{Perk.}} = 1$. \square

Nun betrachten wir die Spinperkolation in der Plus-Phase. Russell Lyons hat dies in [7] in der Ising-Formulierung des Modells getan und dort eine kritische Temperatur berechnet.

Wir geben zunächst die kritische Interaktionsstärke für Spinperkolation in der Plus-Phase an, die sich leicht in eine kritische Temperatur umrechnen läßt.

Satz 19. *Sei (T, \mathbf{P}^+) die Spinperkolation in der Plus-Phase des Broadcasting-Modells mit Interaktionsstärke α . Dann gilt:*

$$\mathbf{P}^+[\hat{T}_r = \infty] = 0, \quad \text{falls } \alpha < \alpha_{c, \text{Perk.}}$$

$$\mathbf{P}^+[\hat{T}_r = \infty] = 1, \quad \text{falls } \alpha > \alpha_{c, \text{Perk.}}$$

Dabei ist $\alpha_{c, \text{Perk.}}$ die kritische Interaktionsstärke für freie Spinperkolation, wie sie in Korollar 8 angegeben ist.

Bevor wir diesen Satz beweisen, geben wir noch in einem Korollar die kritische Temperatur für Spinperkolation in der Plus-Phase an. Wir erhalten sie wieder, dank der eindeutigen Beziehung zwischen Interaktionsstärke und Temperatur, durch einfaches Einsetzen:

Korollar 9. *Im Ising-Modell auf T ist die kritische Temperatur $t_{c, \text{Perk.}}$ für Spinperkolation in der Plus-Phase:*

$$t_{c, \text{Perk.}} = \begin{cases} \left(\text{arccoth} \left(\frac{\text{brT}}{2 - \text{brT}} \right) \right)^{-1} & \text{falls } \text{brT} < 2 \\ +\infty & \text{falls } \text{brT} \geq 2 \end{cases}$$

Beweis. Sei zunächst $\text{brT} < 2$. Wir benutzen den Zusammenhang aus Gleichung (4.18) und die Darstellung des arcoth aus (4.36).

$$\begin{aligned}
t_{c,Perk.} &= t(\alpha_{c,Perk.}) \\
&= \frac{2}{\ln \left[\frac{1}{\text{brT}-1} \right]} \\
&= \left(\frac{1}{2} \ln \left[\frac{2}{2-\text{brT}} \cdot \frac{2-\text{brT}}{2\text{brT}-2} \right] \right)^{-1} \\
&= \left(\frac{1}{2} \ln \left[\left(\frac{\text{brT}}{2-\text{brT}} + 1 \right) / \left(\frac{\text{brT}}{2-\text{brT}} - 1 \right) \right] \right)^{-1} \\
&= \left(\text{arcoth} \left[\frac{\text{brT}}{2-\text{brT}} + 1 \right] \right)^{-1}
\end{aligned}$$

Im Fall $\text{brT} \geq 2$ gilt

$$\begin{aligned}
t_{c,Perk.} &= \lim_{\alpha \rightarrow 1} t(\alpha) \\
&= \lim_{\alpha \rightarrow 1} \frac{2}{\ln [\alpha]} \\
&= \infty
\end{aligned}$$

□

Lyons Beweis dieser Aussage ist deutlich komplexer als unser Beweis von Satz 19. Im Anhang befindet sich eine Variante von Lyons Beweis für den Fall einer kritischen Interaktionsstärke für Spinperkolation in der Plus-Phase des Ising-Modells.

Wir verwenden nun aber die bisherigen Erkenntnisse über das Broadcasting-Modell um einen einfacheren Beweis zu führen:

Beweis von Satz 19. Wir können das Ereignis “*Spinperkolation findet statt*” in die beiden disjunkten Ereignisse “*Spinperkolation findet mit Spin +1 statt*” und “*Spinperkolation findet mit Spin -1 statt*” aufteilen. Wir schreiben dies als:

$$\begin{aligned} \{|\hat{T}_r| = \infty\} &= \{|\hat{T}_r| = \infty, u_r = +1\} \cup \{|\hat{T}_r| = \infty, u_r = -1\} \\ &=: \{|\hat{T}_r^+| = \infty\} \cup \{|\hat{T}_r^-| = \infty\} \end{aligned}$$

Aus den bereits erwähnten Eigenschaften des Broadcasting-Modells, insbesondere aus Satz 7, folgt, daß unendliche Pfade mit Spin +1 in der Plus-Phase mindestens ebenso wahrscheinlich sind wie unendliche Pfade mit mit Spin -1. Also ist:

$$\mathbf{P}^+ [|\hat{T}_r^+| = \infty] \geq \mathbf{P}^+ [|\hat{T}_r^-| = \infty] \quad (4.40)$$

Sei nun $\alpha > \alpha_{c,Perk.}$. Die Abbildung

$$\mathbf{1}_{\{|\hat{T}_r^+| = \infty\}}$$

ist im Sinne von Definition 14 monoton wachsend. denn ein zusätzlicher +1-Spin erhöht die Wahrscheinlichkeit des Ereignisses allenfalls.

\mathbf{P}^+ ist bezüglich der in Definition 15 auf \mathcal{G}_V erklärten Halbordnung das maximale Element, also auch größer als das zum freien Modell gehörige Wahrscheinlichkeitsmaß \mathbf{P} . Die definierende Eigenschaft der Halbordnung erlaubt die Abschätzung von (4.41) nach (4.42):

$$\begin{aligned} \mathbf{P}^+ [|\hat{T}_r| = \infty] &\geq \mathbf{P}^+ [|\hat{T}_r^+| = \infty] \\ &= \int \mathbf{1}_{\{|\hat{T}_r^+| = \infty\}} d\mathbf{P}^+ \end{aligned} \quad (4.41)$$

$$\begin{aligned} &\geq \int \mathbf{1}_{\{|\hat{T}_r^+| = \infty\}} d\mathbf{P} \\ &= \mathbf{P} [|\hat{T}_r^+| = \infty] \end{aligned} \quad (4.42)$$

Die letzte Wahrscheinlichkeit ist echt positiv, weil für $\alpha > \alpha_{c,Perk.}$ in der freien Spinperkolation Perkolation stattfindet, und das freie Broadcasting-Modell keinen Spin in der Wurzel bevorzugt.

Das Ereignis $\{|\hat{T}_r| = \infty\}$ liegt im Tail-Field der u_σ . Für die Plus-Phase enthält dieses Tail-Field, nur Ereignisse, die mit Wahrscheinlichkeit 0 oder 1 eintreten (vgl. [7]). Damit findet für $\alpha > \alpha_{c,Perk.}$ Spinperkolation \mathbf{P}^+ -fast sicher statt.

Sei nun $\alpha < \alpha_{c,Perk.}$. Damit ist nach Satz 18 auch $\alpha < \alpha_{c,P}$. Bei solchen Interaktionsstärken findet kein Phasenübergang statt, es gibt also nur ein Wahrscheinlichkeitsmaß in \mathcal{G}_V . Demnach sind \mathbf{P}^+ und \mathbf{P} identisch. Mit Korollar 8 folgt nun

$$\mathbf{P}^+[\hat{T}_r = \infty] = \mathbf{P}[\hat{T}_r = \infty] = 0 \quad (4.43)$$

□

Der Vergleich von Korollar 8 und Satz 19 zeigt, daß die Randbedingung der Plus-Phase keinen Einfluß auf den kritischen Parameter für Spinperkolation ausübt. Dies ist nicht trivial, denn Spinperkolation in der Plus-Phase unterscheidet sich durchaus von der freien Spinperkolation. Dies zeigt folgendes Beispiel:

Gegeben sei das Broadcasting-Modell mit Fehlerwahrscheinlichkeit ϵ auf einem binären Baum der Höhe n . Die freie Spinperkolation ist hier eine Bernoulli-Perkolation mit Überlebenswahrscheinlichkeit $1 - \epsilon$. Jede Kante wird unabhängig von den anderen Kanten entfernt.

Die Vorgabe der +1-Randbedingung, also $T^n \equiv +1$, zerstört diese Unabhängigkeit. So müssen hier die Zufallsvariablen der zwei Kanten, die von den Knoten unmittelbar vor der n -Sphäre zum Rand laufen, die gleiche Belegung haben. Für $\vec{\sigma} = \{\tau_1, \tau_2\}$ gilt:

$$\mathbf{P}^{n,+}[(\eta_{\tau_1} = +1) \cap (\eta_{\tau_2} = +1)] = 1 - \epsilon \neq \mathbf{P}^{n,+}[\eta_{\tau_1} = +1] \mathbf{P}^{n,+}[\eta_{\tau_2} = +1]$$

$$\mathbf{P}^{n,+}[(\eta_{\tau_1} = -1) \cap (\eta_{\tau_2} = -1)] = \epsilon \neq \mathbf{P}^{n,+}[\eta_{\tau_1} = -1] \mathbf{P}^{n,+}[\eta_{\tau_2} = -1]$$

Spinperkolation in der Plus-Phase ist, anders als die freie Spinperkolation, *keine* Bernoulli-Perkolation. Trotzdem sind die kritischen Interaktionsstärken für beide Perkolationen gleich.

4.8 Zusammenfassung der Ergebnisse

Wir haben in diesem Kapitel das Broadcasting-Modell auf einem Baum T eingeführt und gesehen, daß es auch eine Darstellung als Ising-Modell hat.

Wir haben zwei Arten von Einflußnahme der Spins am Rand des Baums auf die Wurzel charakterisiert. Zum einen haben wir danach gefragt, wann im Broadcasting-Modell Information zwischen dem Spin der Wurzel und den Spins am Rand des Baumes übertragen wird. Wegen Lemma 2 ist dies auch die Frage danach, wann die Belegung in der Wurzel des Baumes und die Belegungen am Rand des Baumes stochastisch unabhängig sind.

Wir konnten zeigen, daß es einen kritischen Interaktionsparameter $\alpha_{c,I}$ für Informationsfluß im Broadcasting-Modell gibt. Für Interaktionen, die geringer als dieser kritische Wert sind, wird keine Information übertragen, oberhalb der kritischen Interaktionsstärke ist die Information zwischen Spin der Wurzel und den Spins am Rand echt positiv.

Die zweite Art der Einflußnahme ist durch die Darstellung als Ising-Modell motiviert. Wir haben uns dabei gefragt, ob die Wahrscheinlichkeit eines Spins in der Wurzel in der Plus-Phase eine andere ist als unter freien Randbedingungen. Wir haben dies als Einfluß der Randbedingung “Rand $\equiv +1$ ” auf den Spin der Wurzel interpretiert.

Auch für dieses Phänomen konnten wir einen kritischen Parameter $\alpha_{c,P}$ angeben. Für Interaktionen, die geringer waren als der kritische Wert, hatten die Spins in der Wurzel in der Plus-Phase gleiche Wahrscheinlichkeit wie unter freien Randbedingungen. Oberhalb des kritischen Wertes war die Wahrscheinlichkeit eines bestimmten Spins der Wurzel in der Plus-Phase $\neq \frac{1}{2}$. Wir haben dies so interpretiert, daß die extreme Randbedingung “Rand $\equiv +1$ ” Einfluß auf die Wurzel ausübt.

Schließlich haben wir die kritischen Parameter $\alpha_{c,P}$ und $\alpha_{c,I}$ miteinander verglichen und gezeigt, daß $\alpha_{c,P}$ stets kleiner als $\alpha_{c,I}$ ist. Wir haben dies so gedeutet, daß der Einfluß der extremen Randbedingungen der Plus- und Minus-Phase auf die Wurzel noch nicht genügen, um Information zwischen Wurzel und Rand zu übertragen. Dies wird anschaulich klar, wenn man bedenkt, daß die Plus- und Minus-Phase als Randbedingungen so exotisch sind, daß sie keinerlei Gewicht im freien Modell erhalten. Daher können sie nichts zum Informationsfluß zwischen Wurzel und Rand beitragen.

Schließlich haben wir das Phänomen der Spinperkolation im Broadcasting-Modell untersucht. Wir haben uns gefragt, ob es Pfade gibt, die in der Wurzel starten, bis zum Rand führen und dabei nur Knoten gleichen Spins besuchen. Wir haben insbesondere untersucht, ob die Wahrscheinlichkeit der Existenz eines solchen Pfades in der Plus-Phase eine andere ist als im freien Modell. Wir konnten zeigen, daß es sowohl unter freien Randbedingungen als auch in der Plus-Phase einen kritischen Parameter $\alpha_{c,Perk.}$ für Spinperkolation gibt. Unterhalb des kritischen Wertes haben Spinperkolationen die Wahrscheinlichkeit 0, während man oberhalb der kritischen Interaktion **P**-fast-sicher Spinperkolation im Broadcasting-Modell findet. Dies bedeutet, daß die Randbedingungen keinen Einfluß auf die Existenz von Spinperkolationen ausüben.

Alle kritischen Parameter hängen lediglich von der Branching-Number des Baumes ab.

Kapitel 5

Weitreichende Einflüsse

Wir reichen in diesem Kapitel die noch ausstehenden Beweise der Sätze 12 und 15 nach. Diese sind zum Teil etwas umfangreicher und benötigen einige Lemmata.

5.1 Eine kritische Interaktionsstärke für Informationsfluß

Wir werden in diesem Abschnitt den noch fehlenden Beweis von Satz 12 führen, der die kritische Interaktionsstärke für Informationsfluß benennt. Dazu beweisen wir zwei Sätze, die den Informationsfluß zwischen Spin der Wurzel und den Spins am Rand des Baumes nach oben und nach unten abschätzen. Die Kombination dieser Aussagen impliziert die Aussage über die kritische Interaktionsstärke.

5.1.1 Information und effektive Leitfähigkeit

Der folgende Satz stellt einen Zusammenhang zwischen dem Broadcasting-Modell und einem elektrischen Netzwerk auf T her. Er zeigt, daß die Information zwischen Spin der Wurzel und den Spins auf einem Cutset Π mit Hilfe des effektiven Widerstandes, wie er in Definition 4 eingeführt wurde, abgeschätzt werden kann.

Satz 20. Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Broadcasting-Modell mit Interaktionsstärke α (bzw. mit Fehlerwahrscheinlichkeit ϵ) auf T und sei $\theta := \frac{\alpha-1}{\alpha+1}$ (bzw. $\theta = 1-2\epsilon$). Jeder Kante $e \in T$ werde eine Leitfähigkeit

$$C(e) = \frac{\theta^{2|e|}}{(1-\theta^2)}$$

zugewiesen. Dann gilt für jeden Cutset Π :

$$I[u_r, u_\Pi] \geq \frac{1}{2} \cdot \frac{1}{1 + R_{eff}(r, \Pi)} \quad (5.1)$$

Der Beweis der Aussage basiert auf der Beobachtung, daß es möglich ist, einen Schätzer S_μ zu konstruieren, der unter Kenntnis der Spins im Cutset Π mit einer Wahrscheinlichkeit $> \frac{1}{2}$ den Spin der Wurzel richtig schätzt. Wenn dies möglich ist, so wird Information zwischen Wurzel und Π übertragen. Der Schätzer S_μ wird dabei vermöge eines geeigneten Flusses μ definiert, was die Brücke zu den elektrischen Netzwerken schlägt.

Wir führen zunächst den schon erwähnten Schätzer S_μ ein und zeigen einige seiner Eigenschaften auf.

Lemma 4. Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Broadcasting-Modell auf T und (T, C) das in Satz 20 beschriebene elektrische Netzwerk. Weiterhin sei Π ein Cutset in T und μ ein Einheitsfluß von r nach Π . Wir definieren:

$$S_\mu := \sum_{\sigma \in \Pi} \frac{\mu(\sigma) u_\sigma}{\theta^{|\sigma|}} \quad (5.2)$$

Dann gilt

$$\mathbf{E}[S_\mu | u_r] = u_r \quad (5.3)$$

$$\mathbf{E}[S_\mu^2] = \mathbf{E}[S_\mu^2 | u_r] = 1 + \sum_{e \in T} R(e) \mu(e)^2 \quad (5.4)$$

Es gilt insbesondere

$$\min_{\mu} \mathbf{E}[S_\mu^2] = 1 + R_{eff}(r, \Pi) \quad (5.5)$$

Das Minimum wird angenommen, wenn μ der **unit current flow** von r nach Π ist.

Beweis. Wir erinnern noch einmal kurz an die Darstellung der u_σ im Broadcasting-System. Jeder Kante e wird eine Zufallsvariable η_e zugeordnet. Dabei gilt:

$$\mathbf{E}[\eta_e] = 1 - \epsilon - \epsilon = \theta \quad (5.6)$$

Die u_σ sind dann:

$$u_\sigma = u_r \prod_{e \in \text{Pfad}(r, \sigma)} \eta_e \quad (5.7)$$

Wir berechnen zunächst $\mathbf{E}[u_\sigma | u_r]$, wobei wir die Darstellung aus (5.7) verwenden:

$$\begin{aligned} \mathbf{E}[u_\sigma | u_r] &= \mathbf{E} \left[u_r \prod_{e \in \text{Pfad}(r, \sigma)} \eta_e \mid u_r \right] \\ &= u_r \prod_{e \in \text{Pfad}(r, \sigma)} \mathbf{E}[\eta_e] \\ &= u_r \theta^{|\sigma|} \end{aligned}$$

Wir benutzen dieses Ergebnis und die Eigenschaft, daß μ ein Einheitsfluß, und damit $\sum_{\sigma \in \Pi} \mu(\sigma) = 1$ ist, und erhalten:

$$\begin{aligned} \mathbf{E}[S_\mu | u_r] &= \mathbf{E} \left[\sum_{\sigma \in \Pi} \frac{\mu(\sigma) u_\sigma}{\theta^{|\sigma|}} \mid u_r \right] \\ &= \sum_{\sigma \in \Pi} \frac{\mu(\sigma)}{\theta^{|\sigma|}} \mathbf{E}[u_\sigma | u_r] \\ &= u_r \sum_{\sigma \in \Pi} \mu(\sigma) \\ &= u_r \end{aligned}$$

Unter Verwendung der Produktdarstellung aus (5.7) erhalten wir:

$$\begin{aligned}
\mathbf{E}[u_\sigma u_\tau] &= \mathbf{E} \left[u_r^2 \cdot \left(\prod_{e \in \text{Pfad}(r, \sigma)} \eta_e \right) \cdot \left(\prod_{e \in \text{Pfad}(r, \tau)} \eta_e \right) \right] \\
&= \mathbf{E} \left[u_r^2 \cdot \left(\prod_{e \in \text{Pfad}(r, \sigma \wedge \tau)} \eta_e^2 \right) \left(\prod_{e \in \text{Pfad}(\sigma \wedge \tau, \sigma)} \eta_e \right) \left(\prod_{e \in \text{Pfad}(\sigma \wedge \tau, \tau)} \eta_e \right) \right] \\
&= \mathbf{E} \left[\prod_{e \in \text{Pfad}(\sigma, \tau)} \eta_e \right] \\
&= \frac{\theta^{|\sigma|} \theta^{|\tau|}}{\theta^{2|\sigma \wedge \tau|}}
\end{aligned}$$

In der Umformung von (5.8) nach (5.9) werden wir folgende Darstellung benutzen:

$$\begin{aligned}
1 + \sum_{e \in \text{Pfad}(\tau)} R(e) &= 1 + \sum_{e \in \text{Pfad}(\tau)} \theta^{-2|e|} - \theta^{-2|e|} \cdot \theta^2 \\
&= 1 + \sum_{e \in \text{Pfad}(\tau)} \theta^{-2|e|} - \theta^{-2(|e|-1)} \\
&= 1 + \theta^{-2|\tau|} - 1 \\
&= \theta^{-2|\tau|}
\end{aligned}$$

Nun erhalten wir für den Erwartungswert von S_μ^2 :

$$\begin{aligned}
\mathbf{E}[S_\mu^2] &= \mathbf{E} \left[\left(\sum_{\sigma \in \Pi} \frac{\mu(\sigma) u_\sigma}{\theta^{|\sigma|}} \right)^2 \right] \\
&= \mathbf{E} \left[\sum_{\sigma, \tau \in \Pi} \frac{\mu(\sigma) \mu(\tau) u_\sigma u_\tau}{\theta^{|\sigma|} \theta^{|\tau|}} \right] \\
&= \sum_{\sigma, \tau \in \Pi} \frac{\mu(\sigma) \mu(\tau)}{\theta^{|\sigma|} \theta^{|\tau|}} \mathbf{E}[u_\sigma u_\tau] \\
&= \sum_{\sigma, \tau \in \Pi} \frac{\mu(\sigma) \mu(\tau)}{\theta^{2|\sigma \wedge \tau|}} \tag{5.8}
\end{aligned}$$

$$= \sum_{\sigma, \tau \in \Pi} \mu(\sigma) \mu(\tau) \cdot \left(1 + \sum_{e \in \text{Pfad}(\sigma \wedge \tau)} R(e) \right) \tag{5.9}$$

$$\begin{aligned}
&= \sum_{\sigma, \tau \in \Pi} \mu(\sigma) \mu(\tau) + \sum_{\substack{\sigma, \tau \in \Pi \\ e \in \text{Pfad}(\sigma \wedge \tau)}} \mu(\sigma) \mu(\tau) R(e) \\
&= 1 + \sum_{e \in T} \left(R(e) \cdot \sum_{\sigma, \tau \in \Pi} \mathbf{1}_{\{e \in \text{Pfad}(\sigma \wedge \tau)\}} \mu(\sigma) \mu(\tau) \right) \tag{5.10}
\end{aligned}$$

Da $\text{Pfad}(\sigma \wedge \tau) = \text{Pfad}(\sigma) \cap \text{Pfad}(\tau)$ ist, gilt:

$$\begin{aligned}
&\sum_{\sigma, \tau \in \Pi} \mathbf{1}_{\{e \in \text{Pfad}(\sigma \wedge \tau)\}} \mu(\sigma) \mu(\tau) \\
&= \left(\sum_{\sigma \in \Pi} \mathbf{1}_{\{e \in \text{Pfad}(\sigma)\}} \mu(\sigma) \right) \left(\sum_{\tau \in \Pi} \mathbf{1}_{\{e \in \text{Pfad}(\tau)\}} \mu(\tau) \right) \\
&= \mu(e) \mu(e) \\
&= \mu(e)^2
\end{aligned}$$

Man erinnere sich daran, daß μ ein Fluß von r nach Π ist und damit $\mu(e) = 0$

ist, falls e außerhalb von Π liegt. Mit der obigen Umformung wird (5.10) zu:

$$\mathbf{E}[S_\mu^2] = 1 + \sum_{e \in \mathbb{T}} R(e) \mu(e)^2 = 1 + E(\mu) \quad (5.11)$$

wobei $E(\mu)$ die in Definition 6 definierte Energie von μ ist. Das Thompson-Prinzip aus Satz 3 besagt, daß die rechte Seite von (5.11) vom **unit current flow** minimiert wird. \square

Beweis von Satz 20. Wir erwarten, daß Information zwischen u_r und u_Π übertragen wird, wenn es möglich ist, einen Schätzer S_μ zu konstruieren, der u_r anhand der Belegung in u_Π mit einer Wahrscheinlichkeit $> \frac{1}{2}$ schätzt. Die Verbindung zwischen S_μ und der Information $I[u_r, u_\Pi]$ schafft nun Lemma 3, das die Information mit dem D_χ -Abstand in Verbindung bringt.

Wir setzen $X := S_\mu(\mathcal{U}^\Pi)$. Damit werden ν_+, ν_- und ν aus (3.4), (3.5) und (3.6) zu:

$$\nu_+(x) = \mathbf{P}[S_\mu^{-1}(x) \mid u_r = +1] \quad , \quad \nu_-(x) = \mathbf{P}[S_\mu^{-1}(x) \mid u_r = -1]$$

$$\nu(x) = \frac{1}{2}\nu_+(x) + \frac{1}{2}\nu_-(x) = \mathbf{P}[S_\mu^{-1}(x)]$$

Mit diesen Bezeichnungen erhalten wir aus der Kombination von Lemma 4 und Lemma 3:

$$I[u_r, u_\Pi] = D_I(\nu_+, \nu_-) \quad (5.12)$$

$$\geq \frac{1}{2} \cdot D_\chi^2(\nu_+, \nu_-) \quad (5.13)$$

$$\geq \frac{\left(\sum_{x \in X} x(\nu_+(x) - \nu_-(x)) \right)^2}{8 \sum_{x \in X} x^2 \nu(x)} \quad (5.14)$$

$$= \frac{(\mathbf{E}[S_\mu \mid u_r = +1] - \mathbf{E}[S_\mu \mid u_r = -1])^2}{8\mathbf{E}[S_\mu^2]}$$

$$= \frac{1}{2} \cdot \frac{(1 - (-1))^2}{4\mathbf{E}[S_\mu^2]}$$

$$\geq \frac{1}{2} \cdot \frac{1}{1 + R_{\text{eff}}(r, \Pi)}$$

In den Umformungen von (5.12) zu (5.13) und (5.13) zu (5.14) haben wir die Ungleichungen aus Lemma 3 benutzt, die einerseits D_I und D_χ^2 miteinander vergleichen und andererseits zeigen, daß

$$(D_\chi(\nu_+, \nu_-))^2 \geq \frac{\left(\sum_{x \in X} x(\nu_+(x) - \nu_-(x)) \right)^2}{4 \sum_{x \in X} x^2 \nu(x)}$$

ist. □

Wir können eine untere Schranke für den Informationsfluß zwischen Wurzel und einem Cutset angeben. Die untere Schranke wird dabei durch den effektiven Widerstand in einem elektrischen Netzwerk bestimmt, dessen Kanten geeignete, von α abhängige Widerstände haben. Man wird erwarten, daß der effektive Widerstand dieses Netzwerkes neben α wesentlich von der Branching-Number des Baumes abhängt. Wir zeigen dies im Beweis von Satz 12.

5.1.2 Unabhängige Version und Dominanz

Das folgende Lemma liefert eine obere Abschätzung für die Information im Broadcasting-Modell. Wir werden diese Abschätzung verwenden, um die kritische Interaktion für Informationsfluß zu berechnen.

Lemma 5. *Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Broadcasting-Modell auf T mit Fehlerwahrscheinlichkeit ϵ (bzw. Interaktionsstärke α).*

Dann gilt für jeden Cutset $\Pi \subset (< T^n)$:

$$I[u_r, u_{T^n}] \leq \sum_{\sigma \in \Pi} I[u_r, u_\sigma] \leq \sum_{\sigma \in \Pi} \theta^{2|\sigma|} \quad (5.15)$$

Dabei ist $\theta = 1 - 2\epsilon$ (bzw. $\theta = \frac{\alpha-1}{\alpha+1}$).

Um diese Aussage zu beweisen, betrachten wir eine verallgemeinerte Version des Broadcasting-Modells. Bisher hatten die den Kanten zugeordneten Zufallsvariablen η_e alle die gleiche Verteilung. In diesem Abschnitt heben wir diese Einschränkung auf.

Sei dazu $(\eta_e), e \in E_T$ eine Familie unabhängiger, aber nicht notwendigerweise identisch verteilter Zufallsvariablen mit Werten in $\{-1, +1\}$ und

$$\mathbf{P}[\eta_e = -1] = \epsilon_e$$

Wie bisher gilt für die Spins der Knoten:

$$u_\sigma = u_r \cdot \prod_{e \in \text{Pfad}(r, \sigma)} \eta_e \quad (5.16)$$

Außerdem setzen wir wie immer:

$$\theta_e := \mathbf{E}[\eta_e] = 1 - 2\epsilon_e \quad (5.17)$$

Wir bezeichnen das auf diesen Zufallsvariablen basierende Modell als **verallgemeinertes Broadcasting-Modell** auf T .

In den nächsten Sätzen seien die erwähnten Broadcasting-Modelle stets solche verallgemeinerten Versionen. Desweiteren meinen wir, wenn wir von einem Baum sprechen, immer einen Baum mit einer den Kanten zugeordneten Familie von Zufallsvariablen $\eta_e, e \in E_T$, die durch die Flipwahrscheinlichkeiten der Kanten festgelegt sind.

Eine Schwierigkeit in der Berechnung von $I[u_r, u_{T^n}]$ besteht darin, daß sich die Knoten der n -Sphäre nicht nur über die Wurzel beeinflussen. Bedingt auf den Zustand der Wurzel, sind die Spins nicht unabhängig, sie können auch über andere Knoten interagieren. Wir werden nun zeigen, daß sich diese Information jedoch durch die Information im Broadcasting-Modell auf einem einfacher handhabbaren Baum abschätzen läßt.

Zunächst sagen wir, was wir unter einer unabhängigen Version eines Baumes verstehen wollen.

Definition 24. *Sei T ein **endlicher** Baum. Jeder Kante e von T sei eine $\{-1, +1\}$ -wertige Zufallsvariable η_e zugeordnet. Wir sagen dann, ein endlicher Baum \tilde{T} mit Zufallsvariablen $\eta_{\tilde{e}}$ ist eine **unabhängige Version** von T , wenn folgendes gilt:*

1. *Für alle $\tilde{\sigma}, \tilde{\tau} \in \partial \tilde{T}$ gilt: $\tilde{\sigma} \wedge \tilde{\tau} = \tilde{r}$.*
2. *Für jeden Ast in T gibt es genau einen Ast in \tilde{T} , wobei die den Kanten zugeordneten $\eta_{\tilde{e}}$ so verteilt sind wie die η_e .*

Offensichtlich sind die Spins der Krone von \tilde{T} , bedingt auf den Zustand der Wurzel, unabhängig. Wir werden auch erwarten, daß die Information zwischen u_r und $u_{\partial T}$ kleiner ist als die Information zwischen $u_{\tilde{r}}$ und $u_{\partial \tilde{T}}$, denn die

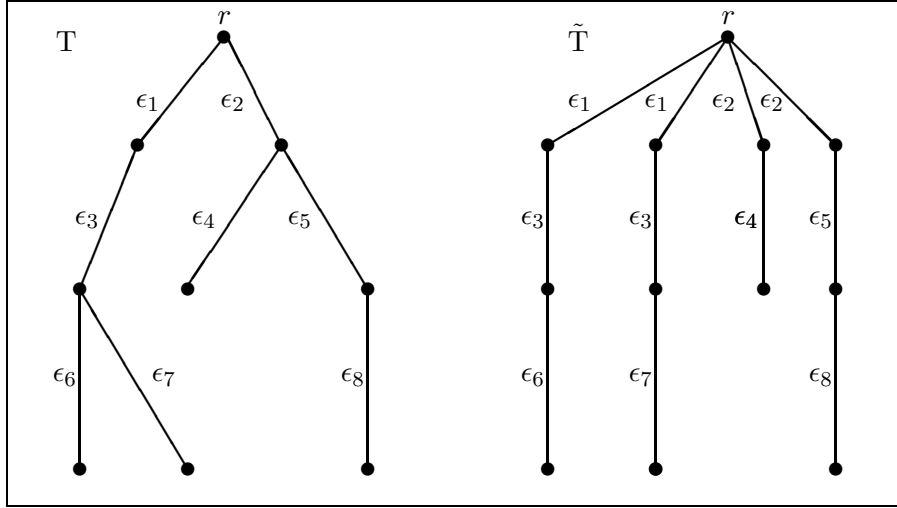


Abbildung 5.1: Ein Baum T und seine unabhängige Version \tilde{T}

zusätzlichen Kanten in \tilde{T} erzeugen eine größere Zufälligkeit in der Verteilung von $u_{\partial\tilde{T}}$. Der Schlüssel zum Beweis dieser Vermutung liegt in der Dominanz des Paares $(u_r, u_{\partial\tilde{T}})$ über das Paar $(u_r, u_{\partial T})$.

Wir werden in Satz 24 nachweisen, daß diese Dominanz tatsächlich gilt. Um den Beweis des Satzes kurz zu halten und um das Beweisprinzip deutlicher zu gestalten, zeigen wir zunächst, daß die gesuchte Eigenschaft in einfachen Spezialfällen gilt.

Im ersten Satz befassen wir uns mit dem einfachsten Fall eines Baumes und seiner unabhängigen Version.

Satz 21. *Seien S und \tilde{S} die Bäume mit den den Kanten zugeordneten Fehlerwahrscheinlichkeiten, wie sie in Abbildung 5.2 gezeigt sind. Dann dominiert das Paar $(u_r, u_{\partial\tilde{S}})$ das Paar $(u_r, u_{\partial S})$.*

Wir haben in der Definition des Dominanz-Begriffs gefordert, daß alle Zufallsvariablen auf dem gleichen Wahrscheinlichkeitsraum definiert sind. Im folgenden denke man dabei an das Produkt der Räume, die durch die Broadcasting-Modelle auf den beiden Bäumen gegeben sind.

Beweis von Satz 21. Wir müssen zeigen, daß es eine Übergangsmatrix von $u_{\partial\tilde{S}}$ nach $u_{\partial S}$ gibt, die die Eigenschaften aus Definition 17 erfüllt. Dazu setzen

wir, wie in Gleichung (5.17) vereinbart:

$$\theta := \mathbf{E}[\eta_e] = 1 - 2\epsilon, \quad \theta_1 := \mathbf{E}[\eta_1] = 1 - 2\epsilon_1, \quad \theta_2 := \mathbf{E}[\eta_{e_2}] = 1 - 2\epsilon_2$$

Außerdem definieren wir

$$q := \frac{1}{2}\left(1 + \frac{\theta_2}{\theta_1}\right)$$

ρ sei eine geeignet zu wählende Wahrscheinlichkeit.

Wenn wir $\mathcal{U}^{\partial S}$ und $\mathcal{U}^{\partial \tilde{S}}$ als $\{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$ anordnen, dann garantiert folgende Übergangsmatrix die gesuchte Dominanz von $(u_r, u_{\partial \tilde{S}})$ über $(u_r, u_{\partial S})$:

$$D := \begin{pmatrix} \rho + (1 - \rho)q & (1 - \rho)(1 - q) & 0 & 0 \\ \rho + (1 - \rho)(1 - q) & (1 - \rho)q & 0 & 0 \\ 0 & 0 & \rho + (1 - \rho)(1 - q) & (1 - \rho)q \\ 0 & 0 & (1 - \rho)(1 - q) & \rho + (1 - \rho)q \end{pmatrix}$$

Bevor wir nun nachweisen, daß eine geeignete Wahl von ρ sicherstellt, daß diese Matrix die in Definition 17 geforderten Eigenschaften erfüllt, wollen wir kurz den zufälligen Mechanismus erklären, der durch sie beschrieben wird.

Dazu sei z eine von den Spins unabhängige Zufallsvariable mit Werten in $\{-1, +1\}$ und $\mathbf{E}[z] = \frac{\theta_2}{\theta_1}$. Damit ist die Verteilung von z festgelegt:

$$\mathbf{P}[z = +1] = \frac{1}{2}\left(1 + \frac{\theta_2}{\theta_1}\right) = q$$

$$\mathbf{P}[z = -1] = \frac{1}{2}\left(1 - \frac{\theta_2}{\theta_1}\right) = 1 - q$$

Wenn nun $b = (b_1, b_2)$ eine Belegung von $u_{\partial \tilde{S}}$ ist, so beschreibt die Übergangsmatrix D , daß mit Wahrscheinlichkeit ρ die Belegung von $u_{\partial S}$ ebenfalls b ist. Mit Wahrscheinlichkeit $1 - \rho$ ist die Belegung $(b_1, z \cdot b_2)$. Wir schreiben dies als:

$$(u_{\tilde{\sigma}_1}, u_{\tilde{\sigma}_2}) \mapsto (u_1^*, u_2^*) := \begin{cases} (u_{\tilde{\sigma}_1}, u_{\tilde{\sigma}_2}) & \text{mit W'keit. } \rho \\ (u_{\tilde{\sigma}_1}, z \cdot u_{\tilde{\sigma}_1}) & \text{mit W'keit. } 1 - \rho \end{cases} \quad (5.18)$$

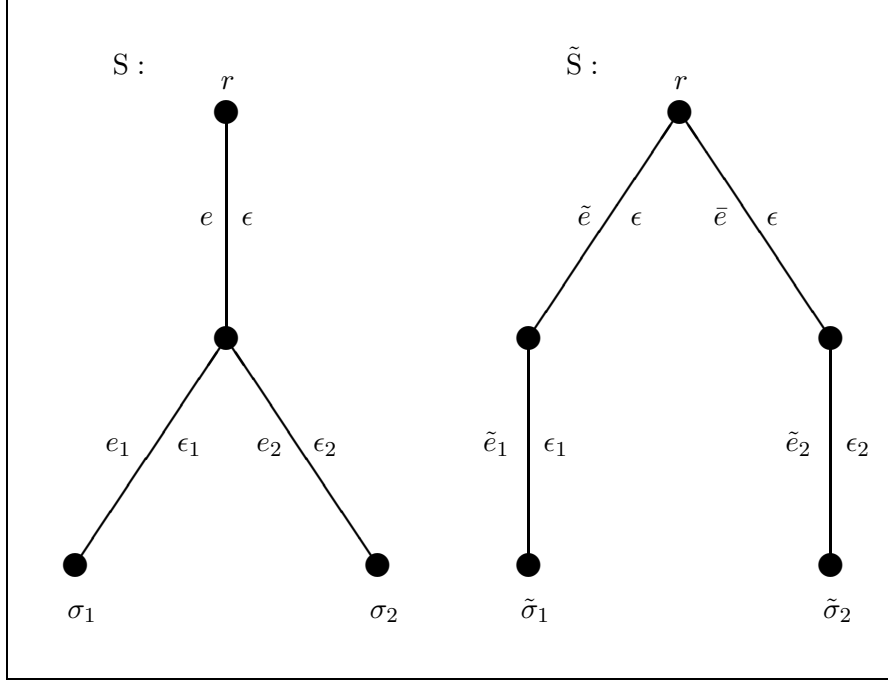


Abbildung 5.2: Die Bäume S und \tilde{S} aus Satz 21

Der Spin von $\tilde{\sigma}_1$ wird stets übernommen, der Spin von $\tilde{\sigma}_2$ wird mit einer gewissen Wahrscheinlichkeit übernommen oder aber durch den Spin in σ_1 ersetzt, wobei vorher noch durch ein zusätzliches Experiment entschieden wird, ob er geflippt wird. Die Wahrscheinlichkeit des Flippens hängt dabei davon ab, ob $u_{\tilde{\sigma}_1}$ gleich $u_{\tilde{\sigma}_2}$ ist oder nicht.

Die Übergangsmatrix gleicht die größere Zufälligkeit in der Verteilung von $u_{\partial\tilde{S}}$ aus. Sie tut dies so, daß die auf den Zustand der Wurzel bedingten Verteilungen von $u_{\partial S}$ und $u_{\partial\tilde{S}}$ übereinstimmen.

Es bleibt noch zu zeigen, daß D die in Definition 17 geforderten Eigenschaften erfüllt. Wir müssen überprüfen, ob für $\zeta \in \{-1, +1\}$ und $b \in \mathcal{U}^{\partial S}$ gilt:

$$\mathbf{P}[u_{\partial S} = b \mid u_r = \zeta] = \sum_{c \in \mathcal{U}^{\partial\tilde{S}}} r(c, b) \cdot \mathbf{P}[u_{\partial\tilde{S}} = c \mid u_r = \zeta]$$

Wir benutzen dazu die Darstellung aus (5.18). Danach müssen wir zeigen, daß für $\zeta \in \{-1, +1\}$ die auf $u_r = \zeta$ bedingten Verteilungen von (u_1^*, u_2^*) und

$(u_{\sigma_1}, u_{\sigma_2})$ übereinstimmen. Es genügt, die Erwartungswerte entsprechender Produkte zu betrachten. Wir verwenden die Produktdarstellung aus (5.16) und erhalten:

$$\mathbf{E}[u_r] = 0$$

$$\mathbf{E}[u_{\sigma_1}] = 0 = \mathbf{E}[u_1^*]$$

$$\mathbf{E}[u_{\sigma_2}] = 0$$

$$= \mathbf{E}[u_2^*]$$

$$= \rho \cdot \mathbf{E}[u_{\tilde{\sigma}_2}] + (1 - \rho) \cdot \mathbf{E}[z] \cdot \mathbf{E}[u_{\tilde{\sigma}_1}]$$

$$= \rho \cdot \mathbf{E}[u_{\tilde{\sigma}_2}] + (1 - \rho) \cdot \frac{\theta_2}{\theta_1} \cdot \mathbf{E}[u_{\tilde{\sigma}_1}] = \mathbf{E}[u_2^*]$$

$$\mathbf{E}[u_r u_{\sigma_1} u_{\sigma_2}] = \mathbf{E}[u_r \cdot u_r \eta_e \eta_{e_1} \cdot u_r \eta_e \eta_{e_2}]$$

$$= \mathbf{E}[u_r^3] \cdot \mathbf{E}[\eta_e^2] \cdot \theta_1 \theta_2$$

$$= 0 \cdot \mathbf{E}[\eta_e^2] \cdot \theta_1 \theta_2$$

$$= 0$$

$$\mathbf{E}[u_r u_1^* u_2^*] = \rho \cdot \mathbf{E}[u_r u_{\tilde{\sigma}_1} u_{\tilde{\sigma}_2}] + (1 - \rho) \cdot \mathbf{E}[u_r u_{\tilde{\sigma}_1} \cdot z \cdot u_{\tilde{\sigma}_1}]$$

$$= \rho \cdot \mathbf{E}[u_r \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1} \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_2}] + (1 - \rho) \cdot \mathbf{E}[u_r \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1} \cdot z \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1}]$$

$$= \mathbf{E}[u_r^3] \cdot (\rho \cdot \mathbf{E}[\eta_{\tilde{e}} \eta_{\tilde{e}} \eta_{\tilde{e}_1} \eta_{\tilde{e}_2}] + (1 - \rho) \cdot \mathbf{E}[\eta_{\tilde{e}}^2 \eta_{\tilde{e}_1}^2 \cdot z])$$

$$= 0$$

$$\mathbf{E}[u_r u_{\sigma_1}] = \mathbf{E}[u_r \cdot u_r \eta_e \eta_{e_1}] = \mathbf{E}[u_r^2] \cdot \theta \theta_1 = \theta \theta_1$$

$$\mathbf{E}[u_r u_1^*] = \mathbf{E}[u_r u_{\tilde{\sigma}_1}] = \mathbf{E}[u_r \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1}] = \mathbf{E}[u_r^2] \cdot \theta \theta_1 = \theta \theta_1$$

$$\mathbf{E}[u_r u_{\sigma_2}] = \theta \theta_2$$

$$\begin{aligned}
\mathbf{E}[u_r u_2^*] &= \rho \cdot \mathbf{E}[u_r u_{\tilde{\sigma}_2}] + (1 - \rho) \cdot \mathbf{E}[u_r \cdot z \cdot u_{\tilde{\sigma}_1}] \\
&= \rho \cdot \theta \theta_2 + (1 - \rho) \cdot \theta \cdot \frac{\theta_2}{\theta_1} \cdot \theta_1 \\
&= \theta \theta_2 \\
\mathbf{E}[u_{\sigma_1} u_{\sigma_2}] &= \mathbf{E}[u_r \eta_e \eta_{e_1} \cdot u_r \eta_e \eta_{e_2}] \\
&= \mathbf{E}[u_r^2] \cdot \mathbf{E}[\eta_e^2] \cdot \theta_1 \theta_2 \\
&= \theta_1 \theta_2
\end{aligned} \tag{5.19}$$

$$\begin{aligned}
\mathbf{E}[u_1^* u_2^*] &= \rho \cdot \mathbf{E}[u_{\tilde{\sigma}_1} u_{\tilde{\sigma}_2}] + (1 - \rho) \cdot \mathbf{E}[u_{\tilde{\sigma}_1} \cdot z \cdot u_{\tilde{\sigma}_1}] \\
&= \rho \cdot \mathbf{E}[u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1} \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_2}] + (1 - \rho) \cdot \mathbf{E}[u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1} \cdot z \cdot u_r \eta_{\tilde{e}} \eta_{\tilde{e}_1}] \\
&= \rho \cdot \theta^2 \theta_1 \theta_2 + (1 - \rho) \cdot \frac{\theta_2}{\theta_1}
\end{aligned} \tag{5.20}$$

Damit auch die beiden Ausdrücke aus Gleichung (5.19) und Gleichung (5.20) übereinstimmen, müssen wir ρ geeignet wählen:

$$\begin{aligned}
\mathbf{E}[u_{\sigma_1} u_{\sigma_2}] &= \mathbf{E}[u_1^* u_2^*] \\
\iff \theta_1 \theta_2 &= \rho \cdot \theta^2 \theta_1 \theta_2 + (1 - \rho) \cdot \frac{\theta_2}{\theta_1} \\
\iff \rho &= \frac{\theta_1^2 - 1}{\theta^2 \theta_1^2 - 1}
\end{aligned} \tag{5.21}$$

Die Matrix D erfüllt mit der Wahl von ρ aus (5.21) die in der Definition der Dominanz geforderten Eigenschaften, also dominiert $(u_r, u_{\partial \tilde{S}})$ das Paar $(u_r, u_{\partial S})$. Insbesondere können wir für jede Wahl von $\epsilon, \epsilon_1, \epsilon_2$ eine solche Übergangsmatrix $D = D_{\epsilon, \epsilon_1, \epsilon_2}$ konstruieren, die die Dominanz gewährleistet. \square

Für den einfachsten Fall einen Baumes und seiner unabhängigen Version haben wir gezeigt, daß das Paar $(u_r, u_{\partial \tilde{S}})$ das Paar $(u_r, u_{\partial S})$ dominiert. Korollar 2, das die Dominanz mit der gegenseitigen Information in Bezug setzt, sagt, daß damit $I[u_r, u_{\partial \tilde{S}}] \geq I[u_r, u_{\partial S}]$ ist. Diese Abschätzung möchten wir für beliebige endliche Bäume nachweisen.

Zunächst wenden wir uns Bäumen zu, die S und \tilde{S} sehr ähnlich sind.

Satz 22. *Seien V und V' die Bäume, die man erhält, wenn man an die Wurzel jeweils den gleichen endlichen Teilbaum heftet. Dann gilt, $(u_r, u_{\partial V'})$ dominiert $(u_r, u_{\partial V})$.*

Beweis. Wir bezeichnen die Kronen der angehefteten Bäume als C und C' , wie dies in Abbildung 5.1.2 geschehen ist.

u_C und $u_{C'}$ sind bedingt auf den Spin der Wurzel unabhängig von den Spins in den anderen Knoten der Krone. Außerdem sind die Verteilungen von u_C und $u_{C'}$ gleich, da die Teilbäume, aus denen sie stammen, identisch sind. Die Einträge der gesuchten Übergangsmatrix hängen also nur den Belegungen außerhalb von C und C' ab.

Wenn wir $\mathcal{U}^C = \mathcal{U}^{C'}$ als $\{c_1, \dots, c_n\}$ anordnen und analog $\mathcal{U}^{\partial V} = \mathcal{U}^{\partial V'}$ als

$$\{(1, 1, c_1), (1, -1, c_1), (-1, 1, c_1), (-1, -1, c_1), (1, 1, c_2), \dots, (-1, -1, c_n)\}$$

anordnen, ist

$$\hat{D} := \hat{D}_{\epsilon, \epsilon_1, \epsilon_2} := \begin{pmatrix} D & 0 & \dots & 0 \\ 0 & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D \end{pmatrix} \quad (5.22)$$

die Übergangsmatrix, die die Dominanz von $(u_r, u_{\partial V'})$ über $(u_r, u_{\partial V})$ garantiert. Dabei ist D die 4×4 -Matrix aus dem Beweis von Satz 21.

\hat{D} erfüllt die gesuchten Eigenschaften, denn die Verteilung von $u_{C'}$ wird nicht verändert, und auf den anderen Knoten werden die Gewichte so umverteilt, wie es die Untermatrix D beschreibt. Von D wissen wir aus Satz 21, daß dies auf die richtige Art geschieht. \square

Ein ähnliche Überlegung wie diese wird uns auch im nächsten Beweis hilfreich sein. Wir werden dort zeigen, daß die entsprechende Dominanz für jeden endlichen Baum T und seine *um einen Pfad unabhängige Version* gilt. Definieren wir zunächst, was wir darunter verstehen wollen:

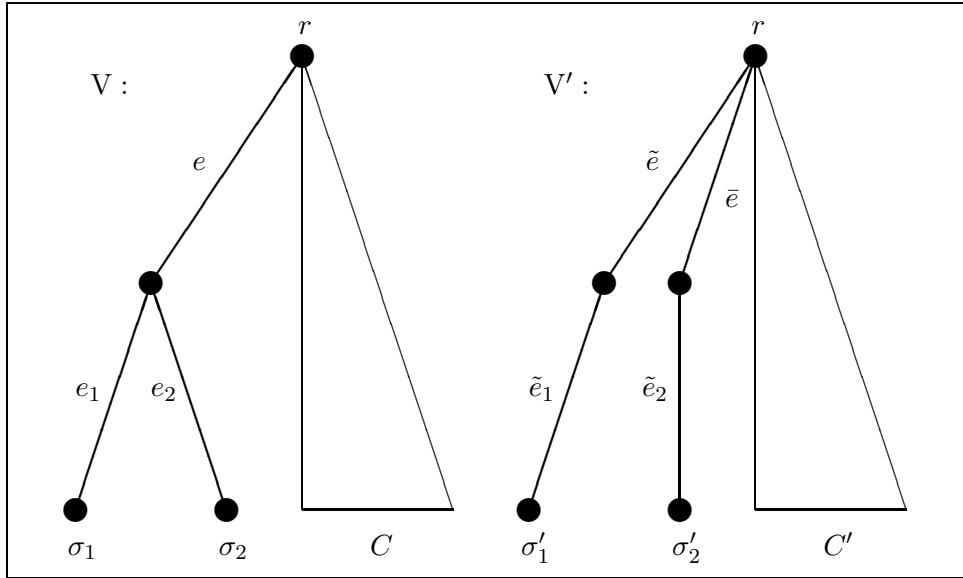


Abbildung 5.3: Die Bäume V und V' aus Satz 22

Definition 25. Sei T ein endlicher Baum, der nicht gleich seiner unabhängigen Version ist. Sei $\tau \in T$ der der Wurzel am nächsten gelegene Knoten, der mehr als einen Nachkommen hat. Gibt es mehrere solcher Knoten, so sei τ ein beliebiger, aber fest gewählter. A seien die Blätter in der Krone von T , die von einem fest gewählten Nachfolger von τ stammen. B seien die Blätter der Krone, die von den anderen Nachfolgern von τ stammen und C seien die Blätter, die nicht von τ abstammen.

Der Baum T' habe in Bezug auf Größe und Tiefe die gleiche Krone wie T . Auch der Teilbaum mit Blättern C' sei identisch zu dem Teilbaum in T , der die Krone C hat. Der Pfad von r zu τ wird in T' durch zwei identische Kopien (in Bezug auf Pfadlänge und Flipwahrscheinlichkeiten der Kanten) von r zu τ' bzw. von r zu $\bar{\tau}$ ersetzt. An τ' wird nun der Teilbaum aus T geheftet, dessen Blätter A sind, an $\bar{\tau}$ wird der Teilbaum mit den Blättern B geheftet.

Dann nennen wir T' eine **um einen Pfad unabhängige Version** von T .

Abbildung 5.4 zeigt einen Baum und seine um einen Pfad unabhängige Version. Ein Beispiel hierfür haben wir bereits in Satz 22 kennengelernt. Dort haben wir auch schon die Dominanz nachgewiesen, die wir nun allgemein feststellen:

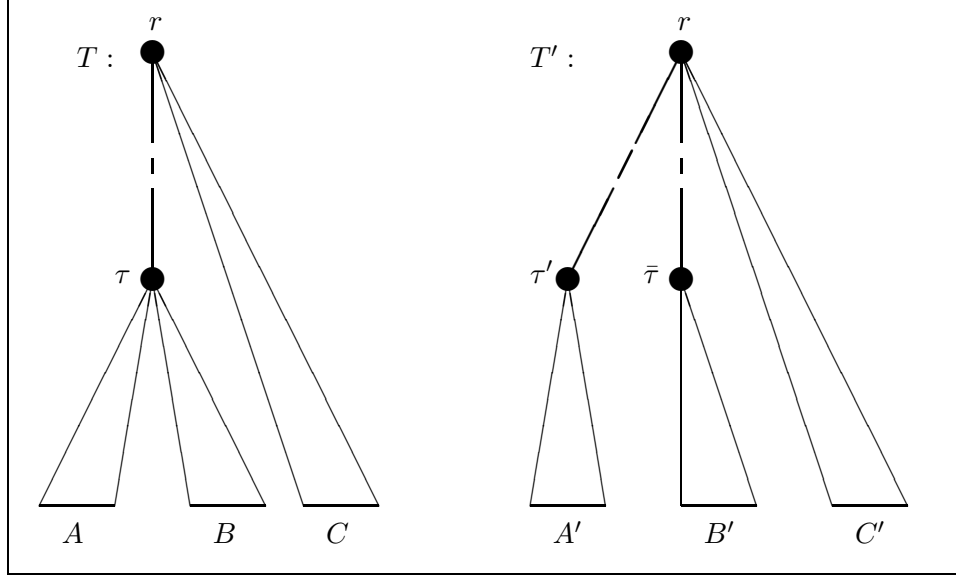


Abbildung 5.4: Ein Baum und seine um einen Pfad unabhängigere Version

Satz 23. *Sei T ein endlicher Baum, der nicht gleich seiner unabhängigen Version ist. T' sei die um einen Pfad unabhängigere Version von T . Dann wird $(u_r, u_{\partial T})$ von $(u_r, u_{\partial T'})$ dominiert.*

Beweis: Bevor wir die Übergangsmatrix von $u_{\partial T'}$ zu $u_{\partial T}$ angeben, wollen wir die Verteilung dieser Zufallsvariablen etwas genauer betrachten. Wir definieren dazu für $\zeta \in \{-1, +1\}$:

$$\epsilon := \mathbf{P}[u_\tau \neq \zeta \mid u_r = \zeta] = \mathbf{P}[u_{\tau'} \neq \zeta \mid u_r = \zeta] = \mathbf{P}[u_{\bar{\tau}} \neq \zeta \mid u_r = \zeta] \quad (5.23)$$

ϵ ist die Wahrscheinlichkeit dafür, daß das Signal der Wurzel in τ geflippt ankommt. Weiterhin definieren wir für $a \in \mathcal{U}^A, b \in \mathcal{U}^B$:

$$p_+(a) := \mathbf{P}[u_A = a \mid u_\tau = +1] = \mathbf{P}[u_{A'} = a \mid u_{\tau'} = +1]$$

$$p_-(a) := \mathbf{P}[u_A = a \mid u_\tau = -1] = \mathbf{P}[u_{A'} = a \mid u_{\tau'} = -1]$$

$$q_+(b) := \mathbf{P}[u_B = b \mid u_\tau = +1] = \mathbf{P}[u_{B'} = b \mid u_{\bar{\tau}} = +1]$$

$$q_-(b) := \mathbf{P}[u_B = b \mid u_\tau = -1] = \mathbf{P}[u_{B'} = b \mid u_{\bar{\tau}} = -1]$$

Mit diesen Bezeichnungen erhalten wir

$$\begin{aligned} \mathbf{P}[u_A = a, u_B = b \mid u_r = -1] \\ = (1 - \epsilon) \cdot p_-(a)q_-(b) + \epsilon \cdot p_+(a)q_+(b) \end{aligned} \quad (5.24)$$

$$\neq ((1 - \epsilon)p_-(a) + \epsilon p_+(a)) \cdot ((1 - \epsilon)q_-(b) + \epsilon q_+(b)) \quad (5.25)$$

$$= \mathbf{P}[u_{A'} = a, u_{B'} = b \mid u_r = -1]$$

Da u_C und $u_{C'}$ bedingt auf den Spin der Wurzel unabhängig von den Spins der anderen Blätter sind und ihre Verteilungen übereinstimmen, gilt für jedes $c \in \mathcal{U}^C$ und $\zeta \in \{-1, +1\}$:

$$\mathbf{P}[u_A = a, u_B = b, u_C = c \mid u_r = \zeta] \neq \mathbf{P}[u_{A'} = a, u_{B'} = b, u_{C'} = c \mid u_r = \zeta]$$

Wie wir nicht anders erwartet haben, sind die auf $u_r = \zeta$ bedingten Verteilungen von $u_{\partial T'}$ und $u_{\partial T}$ nicht gleich. Es zeigt sich aber, daß die bedingten Wahrscheinlichkeiten bestimmter Ereignisse gleich sind. Wir benutzen im folgenden mit $a \in \mathcal{U}^A, b \in \mathcal{U}^B, c \in \mathcal{U}^C$ die abkürzende Schreibweise:

$$\{\pm a, \pm b\} := \{(a, b), (a, -b), (-a, b), (-a, -b)\}$$

$$\{\pm a, \pm b, c\} := \{(a, b, c), (a, -b, c), (-a, b, c), (-a, -b, c)\}$$

Die auf $u_r = \zeta$ bedingten Wahrscheinlichkeiten von $u_{\partial T} \in \{\pm a, \pm b, c\}$ bzw. $u_{\partial T'} \in \{\pm a, \pm b, c\}$ sind gleich.

Dies zeigt die folgende Rechnung, die die Darstellung aus (5.24) verwendet:

$$\mathbf{P}[(u_A, u_B) \in \{\pm a, \pm b\} \mid u_r = -1] \quad (5.26)$$

$$\begin{aligned} &= (1 - \epsilon)p_-(a)q_-(b) + \epsilon p_+(a)q_+(b) \\ &\quad + (1 - \epsilon)p_-(a)q_-(-b) + \epsilon p_+(a)q_+(-b) \\ &\quad + (1 - \epsilon)p_-(-a)q_-(b) + \epsilon p_+(-a)q_+(b) \\ &\quad + (1 - \epsilon)p_-(-a)q_-(-b) + \epsilon p_+(-a)q_+(-b) \\ &= p_-(a)q_-(b) + p_-(a)q_+(b) + p_+(a)q_-(b) + p_+(a)q_+(b) \end{aligned} \quad (5.27)$$

Auf T' erhalten wir mit der Darstellung aus (5.25) in einer ähnlichen Rechnung:

$$\mathbf{P}[(u'_A, u'_B) \in \{\pm a, \pm b\} \mid u_r = -1] \quad (5.28)$$

$$\begin{aligned} &= [(1 - \epsilon)p_-(a) + \epsilon p_+(a)] \cdot [(1 - \epsilon)q_-(b) + \epsilon q_+(b)] \\ &\quad + [(1 - \epsilon)p_-(a) + \epsilon p_+(a)] \cdot [(1 - \epsilon)q_-(-b) + \epsilon q_+(-b)] \\ &\quad + [(1 - \epsilon)p_-(-a) + \epsilon p_+(-a)] \cdot [(1 - \epsilon)q_-(b) + \epsilon q_+(b)] \\ &\quad + [(1 - \epsilon)p_-(-a) + \epsilon p_+(-a)] \cdot [(1 - \epsilon)q_-(-b) + \epsilon q_+(-b)] \\ &= p_-(a)q_-(b) + p_-(a)q_+(b) + p_+(a)q_-(b) + p_+(a)q_+(b) \end{aligned} \quad (5.29)$$

Da u_C und $u_{C'}$ bedingt auf den Spin der Wurzel unabhängig von den Spins der anderen Blätter sind, folgt:

$$\mathbf{P}[u_{\partial T} \in \{\pm a, \pm b, c\} \mid u_r = -1] = \mathbf{P}[u_{\partial T'} \in \{\pm a, \pm b, c\} \mid u_r = -1] \quad (5.30)$$

Dies bedeutet, daß die gesuchte Übergangsmatrix nur die Gewichte auf den Ereignissen in $\{\pm a, \pm b, c\}$ umverteilen muß.

Um die Übergangsmatrix angeben zu können, müssen wir eine Anordnung der Zustände in $\mathcal{U}^{\partial T}$ festlegen. Wir ordnen dazu \mathcal{U}^A als $\{a_1, -a_1, \dots, a_n, -a_n\}$, \mathcal{U}^B als $\{b_1, -b_1, \dots, b_m, -b_m\}$ und \mathcal{U}^C als $\{c_1, \dots, c_k\}$ an. $\mathcal{U}^{\partial T} = \mathcal{U}^{\partial T'}$, das Produkt dieser Mengen, ordnen wir als

$$\mathcal{U}^{\partial T} = \{(a_1, b_1, c_1), (a_1, -b_1, c_1), (-a_1, b_1, c_1), (-a_1, -b_1, c_1), (a_1, b_1, c_2), \dots\}$$

an. $\mathcal{U}^{\partial T'}$ sei ebenso angeordnet.

Bezüglich dieser Anordnung hat die gesuchte Übergangsmatrix die Form:

$$D' = \begin{pmatrix} D_{\pm a_1, \pm b_1} & 0 & \dots & 0 \\ 0 & D_{\pm a_1, \pm b_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D_{\pm a_n, \pm b_m} \end{pmatrix} \quad (5.31)$$

Dabei sind die $D_{\pm a_i, \pm b_i}$ 4×4 -Matrizen, deren Einträge nur von der Belegung auf A und B bzw. A' und B' abhängen.

Wir widmen uns nun dem Aufbau dieser Untermatrizen und werden dabei sehen, daß sie den Matrizen D aus Satz 21 entsprechen, wenn wir ϵ, ϵ_1 und ϵ_2 geeignet wählen.

Seien $a \in \mathcal{U}^A, b \in \mathcal{U}^B, c \in \mathcal{U}^C$ fest gewählt. Wir definieren:

$$\epsilon_1 := \frac{p_+(-a)}{p_+(a) + p_+(-a)}, \quad \theta_1 := 1 - 2\epsilon_1 \quad (5.32)$$

$$\epsilon_2 := \frac{q_+(-b)}{q_+(b) + q_+(-b)}, \quad \theta_2 := 1 - 2\epsilon_2 \quad (5.33)$$

Durch Umbenennung von a zu $-a$ bzw. von b zu $-b$ ist es stets möglich, die Bezeichnungen so zu wählen, daß $0 \leq \epsilon_1 \leq \epsilon_2$ ist. Sei diese Wahl geschehen.

Seien nun V und V' die Bäume aus Abbildung 5.1.2, wobei die Teilbäume mit Krone C bzw. C' die entsprechenden Teilbäume aus T und T' sind. Die Flipwahrscheinlichkeiten an den übrigen Kanten von V und V' seien ϵ aus Gleichung (5.23) sowie ϵ_1 und ϵ_2 aus den Gleichungen (5.32) und (5.33).

Dann sind folgende bedingte Wahrscheinlichkeiten auf V und T gleich:

$$\begin{aligned} & \mathbf{P}[u_{\sigma_1} = 1, u_{\sigma_2} = -1 \mid u_r = 1] \\ &= (1 - \epsilon)(1 - \epsilon_1)\epsilon_2 + \epsilon\epsilon_1(1 - \epsilon_2) \\ &= (1 - \epsilon) \frac{p_+(a)}{p_+(a) + p_+(-a)} \cdot \frac{q_+(-b)}{q_+(b) + q_+(-b)} \\ &\quad + \epsilon \frac{p_+(-a)}{p_+(a) + p_+(-a)} \cdot \frac{q_+(b)}{q_+(b) + q_+(-b)} \\ &= \mathbf{P}[u_A = a, u_B = -b \mid (u_A, u_B) \in \{\pm a, \pm b\}, u_r = 1] \end{aligned}$$

Diese Ereignisse haben auch auf T' und V' die gleichen bedingten Wahrscheinlichkeiten:

$$\begin{aligned}
& \mathbf{P}[u_{\sigma'_1} = 1, u_{\sigma'_2} = -1 \mid u_r = 1] \\
&= ((1 - \epsilon)(1 - \epsilon_1) + \epsilon\epsilon_1) \cdot ((1 - \epsilon)\epsilon_2 + \epsilon(1 - \epsilon_2)) \\
&= \left((1 - \epsilon) \frac{p_+(a)}{p_+(a) + p_+(-a)} + \epsilon \frac{p_+(-a)}{p_+(a) + p_+(-a)} \right) \\
&\quad \cdot \left((1 - \epsilon) \frac{q_+(-b)}{q_+(b) + q_+(-b)} + \epsilon \frac{q_+(b)}{q_+(b) + q_+(-b)} \right) \\
&= \mathbf{P}[u_{A'} = a, u_{B'} = -b \mid (u_{A'}, u_{B'}) \in \{\pm a, \pm b\}, u_r = 1]
\end{aligned}$$

Wie auch zuvor, bedeutet dies für jedes $c \in \mathcal{U}^C$:

$$\begin{aligned}
& \mathbf{P}[u_{\sigma_1} = 1, u_{\sigma_2} = -1, u_C = c \mid u_r = 1] \\
&= \mathbf{P}[u_A = a, u_B = -b, u_C = c \mid u_r = 1, (u_A, u_B, u_C) \in \{\pm a, \pm b, c\}] \\
& \mathbf{P}[u_{\sigma'_1} = 1, u_{\sigma'_2} = -1, u_{C'} = c \mid u_r = 1] \\
&= \mathbf{P}[u_{A'} = a, u_{B'} = -b, u_{C'} = c \mid u_r = 1, (u_{A'}, u_{B'}, u_{C'}) \in \{\pm a, \pm b, c\}]
\end{aligned}$$

Die Gleichheit dieser bedingten Wahrscheinlichkeiten gilt auch für alle anderen Belegungen aus $\{\pm a, \pm b, c\}$ und $\{\pm 1, \pm 1, c\}$, wenn wir die Zuordnungen

$$a \leftrightarrow 1, \quad b \leftrightarrow 1$$

wählen. Dies folgt mit ganz ähnlichen Rechnungen, ist aber auch nach Konstruktion von V und V' klar.

Die Dominanz von $(u_r, u_{\partial V'})$ über $(u_r, u_{\partial V})$ wird durch die Matrix \hat{D} aus (5.22) gewährleistet. Diese Matrix verteilt für $c \in \mathcal{U}^C$ jeweils die Gewichte auf $\{\pm 1, \pm 1, c\}$ um. Die Untermatrix $D_{\pm a, \pm b}$ muß die Gewichte genauso auf $\{\pm a, \pm b, c\}$ umverteilen wie die entsprechende Untermatrix in \hat{D} .

Wir setzen also mit den Werten von ϵ, ϵ_1 und ϵ_2 aus (5.23), (5.32) und (5.33):

$$D_{\pm a, \pm b} := D_{\epsilon, \epsilon_1, \epsilon_2}$$

Wir zeigen nun, daß diese Wahl die Dominanz gewährleistet.

Wegen der Form von \hat{D} aus (5.22) und D' aus (5.31) gilt:

$$\hat{d}(y, (1, 1, c)) \neq 0 \quad \Rightarrow \quad y \in \{\pm 1, \pm 1, c\}$$

$$d'(y, (a, b, c)) \neq 0 \quad \Rightarrow \quad y \in \{\pm a, \pm b, c\}$$

Die oben gezeigte Gleichheit der bedingten Wahrscheinlichkeiten auf V und T gestattet nun folgende Rechnung, bei der $\zeta \in \{-1, +1\}$ ist:

$$\begin{aligned} & \mathbf{P}[u_{\partial T} = (a, b, c) \mid u_r = \zeta, u_{\partial T} \in \{\pm a, \pm b, c\}] \\ &= \mathbf{P}[u_{\partial V} = (1, 1, c) \mid u_r = \zeta] \\ &= \sum_{y \in \mathcal{U}^{\partial V}} \hat{r}(y, (1, 1, c)) \cdot \mathbf{P}[u_{\partial V'} = y \mid u_r = \zeta] \\ &= \sum_{y \in \{\pm 1, \pm 1, c\}} \hat{r}(y, (1, 1, c)) \cdot \mathbf{P}[u_{\partial V'} = y \mid u_r = \zeta] \\ &= \sum_{y \in \{\pm a, \pm b, c\}} r(y, (a, b, c)) \cdot \mathbf{P}[u_{\partial T'} = y \mid u_r = \zeta, u_{\partial T'} \in \{\pm a, \pm b, c\}] \\ &= \sum_{y \in \mathcal{U}^{\partial T'}} r(y, (a, b, c)) \cdot \mathbf{P}[u_{\partial T'} = y \mid u_r = \zeta, u_{\partial T'} \in \{\pm a, \pm b, c\}] \quad (5.34) \end{aligned}$$

Nun ist

$$\begin{aligned} \mathbf{P}[u_{\partial T} = (a, b, c) \mid u_r = \zeta] &= \mathbf{P}[u_{\partial T} = (a, b, c) \mid u_r = \zeta, u_{\partial T'} \in \{\pm a, \pm b, c\}] \\ &\quad \cdot \mathbf{P}[u_{\partial T} \in \{\pm a, \pm b, c\}] \end{aligned}$$

Da nach (5.30) die auf $u_r = \zeta$ bedingten Wahrscheinlichkeiten der Ereignisse $u_{\partial T} \in \{\pm a, \pm b, c\}$ und $u_{\partial T'} \in \{\pm a, \pm b, c\}$ gleich sind, folgt:

$$\begin{aligned}
& \mathbf{P}[u_{\partial T} = (a, b, c) \mid u_r = \zeta] \\
&= \left(\sum_{y \in \mathcal{U}^{\partial T'}} d(y, (a, b, c)) \cdot \mathbf{P}[u_{\partial T'} = y \mid u_r = \zeta, u_{\partial T'} \in \{\pm a, \pm b, c\}] \right) \\
&\quad \cdot \mathbf{P}[u_{\partial T} \in \{\pm a, \pm b, c\} \mid u_r = \zeta] \\
&= \left(\sum_{y \in \mathcal{U}^{\partial T'}} d(y, (a, b, c)) \cdot \mathbf{P}[u_{\partial T'} = y \mid u_r = \zeta, u_{\partial T'} \in \{\pm a, \pm b, c\}] \right) \\
&\quad \cdot \mathbf{P}[u_{\partial T'} \in \{\pm a, \pm b, c\} \mid u_r = \zeta] \\
&= \sum_{y \in \mathcal{U}^{\partial T'}} d(y, (a, b, c)) \cdot \mathbf{P}[u_{\partial T'} = y \mid u_r = \zeta]
\end{aligned}$$

Dies bedeutet, daß $(u_r, u_{\partial T})$ von $(u_r, u_{\partial T'})$ dominiert wird. \square

Wir haben bewiesen, daß die gesuchte Dominanz für jeden endlichen Baum und seine um einen Pfad unabhängige Version gilt. Eine direkte Konsequenz ist der folgende Satz.

Satz 24. *Sei T ein endlicher Baum und \tilde{T} seine unabhängige Version. Dann wird $(u_r, u_{\partial T})$ von $(u_r, u_{\partial \tilde{T}})$ dominiert.*

Beweis. Es gibt Bäume T_1, \dots, T_n mit $T_1 = T$ und $T_n = \tilde{T}$, sodaß T_{i+1} eine um einen Pfad unabhängige Version von T_i ist. Nach Satz 23 gilt die gesuchte Dominanz zwischen T_{i+1} und T_i . Wir nennen die zugehörige Übergangsmatrix D_i . Dann ist

$$\tilde{D} := D_1 \cdot \dots \cdot D_{n-1} \tag{5.35}$$

eine Übergangsmatrix von $u_{\partial \tilde{T}}$ nach $u_{\partial T}$, die die Dominanz von $(u_r, u_{\partial \tilde{T}})$ über $(u_r, u_{\partial T})$ garantiert. \square

Als eine Folgerung aus Satz 24 erhalten wir die gesuchte obere Abschätzung für die Information zwischen Spin der Wurzel und den Spins der Krone auf einem endlichen Baum:

Korollar 10. *Sei T ein **endlicher** Baum und \tilde{T} seine unabhängige Version. Dann gilt für die Spins in den Broadcasting-Modellen auf T und \tilde{T} :*

$$I[u_r, u_{\partial T}] \leq I[u_r, u_{\partial \tilde{T}}] \leq \sum_{\sigma \in \partial \tilde{T}} I[u_r, u_\sigma] \leq \sum_{\sigma \in \partial \tilde{T}} \theta^{2|\sigma|} \quad (5.36)$$

Beweis. Nach Satz 24 dominiert das Paar $(u_r, u_{\partial \tilde{T}})$ das Paar $(u_r, u_{\partial T})$. Diese Dominanz bedeutet mit Korollar 2, daß die Information zwischen u_r und $u_{\tilde{T}}$ größer ist als die Information zwischen u_r und $u_{\partial T}$. Damit ist die erste Ungleichung bewiesen.

In \tilde{T} schneiden sich die Äste nur in der Wurzel. Bedingt auf den Spin der Wurzel sind die $u_{\sigma'}, \sigma' \in T'$ unabhängig. Die zweite Ungleichung ist somit eine Konsequenz aus Lemma 2.

In Satz 11 haben wir gezeigt, daß im Broadcasting-Modell folgende Abschätzung für die Information gilt:

$$I[u_r, u_\sigma] \leq \mathbf{E}[u_r u_\sigma]^2 = \theta^{2|\sigma|}$$

Dies sichert die dritte Ungleichung. □

Nun können wir den noch ausstehenden Beweis von Lemma 5 leicht führen. Wir müssen nur einen geeigneten endlichen Ausschnitt in T , der jetzt wieder ein Baum *ohne* endliche Äste ist, betrachten:

Beweis von Lemma 5. Sei Π ein Cutset in T und $n \in \mathbb{N}$ so gewählt, daß $\Pi \subset \leq T^n$ ist.

Das Broadcasting-Modell ist ein Markov-Feld, und daher ist $u_r \rightarrow u_\Pi \rightarrow u_{T^n}$ eine Markov-Kette. Mit dem Data-processing-Lemma folgt:

$$I[u_r, u_{T^n}] \leq I[u_r, u_\Pi] \quad (5.37)$$

Wir stellen uns nun vor, daß alle Kanten und Knoten, die außerhalb von Π liegen, aus T entfernt werden. Den so entstandenen endlichen Baum nennen wir T_* , seine unabhängige Version \tilde{T}_* .

Aus der Definition der unabhängigen Version folgt, daß die Krone von \tilde{T}_* in Bezug auf Anzahl und Tiefe der Knoten identisch zu Π ist. Damit gilt:

$$\sum_{\sigma \in \Pi} I[u_r, u_\sigma] = \sum_{\sigma \in \partial \tilde{T}_*} I[u_r, u_\sigma]$$

Mit Korollar 10 folgt nun

$$I[u_r, u_\Pi] = I[u_r, u_{\partial T_*}] \leq \sum_{\sigma \in \partial \tilde{T}_*} I[u_r, u_\sigma] = \sum_{\sigma \in \Pi} I[u_r, u_\sigma] \leq \sum_{\sigma \in \partial \tilde{T}_*} \theta^{2|\sigma|} = \sum_{\sigma \in \Pi} \theta^{2|\sigma|}$$

Die Kombination dieser Abschätzung mit (5.37) beweist die Behauptung:

$$I[u_r, u_{T^n}] \leq I[u_r, u_\Pi] \leq \sum_{\sigma \in \Pi} \theta^{2|\sigma|}$$

□

Wir haben in diesem Abschnitt eine obere Abschätzung für die Information zwischen dem Spin der Wurzel und den Spins der Knoten eines Cutset nachgewiesen. Da die n -Sphäre T^n auch ein Cutset ist, haben wir damit auch eine Abschätzung für $I[u_r, u_{T^n}]$.

Die obere Schranke ist eine Cutset-Summe, wie sie auch in der Definition der Branching-Number (Definition 1) auftaucht. Wir können also erwarten, daß sich der kritische Wert für Informationsfluß durch die Branching-Number ausdrücken läßt.

5.1.3 Kritische Interaktion für Informationsfluß

Die Abschätzungen des Informationsflusses nach oben und nach unten ermöglichen es uns, nun die Aussage über die kritische Interaktion zu beweisen. Wir rekapitulieren noch einmal die zu beweisende Aussage:

Satz 13. *Im Broadcasting-Modell mit Interaktionsstärke α gilt für die Information zwischen Spin der Wurzel und den Spins der n -Sphäre:*

$$\lim_{n \rightarrow \infty} I[u_r, u_{T^n}] = 0, \quad \text{falls } \alpha < \alpha_{c,I}$$

$$\lim_{n \rightarrow \infty} I[u_r, u_{T^n}] > 0, \quad \text{falls } \alpha > \alpha_{c,I}$$

wobei

$$\alpha_{c,I} := \frac{\sqrt{\text{brT}} + 1}{\sqrt{\text{brT}} - 1} \quad (5.38)$$

ist.

Beweis. Wir erklären kurz, wieso der Grenzwert von $I[u_r, u_{T^n}]$ für $n \rightarrow \infty$ existiert, bevor wir zeigen, daß $\alpha_{c,I}$ kritisch für Informationsfluß ist.

Die Markov-Feld-Eigenschaft des Broadcasting-Modells garantiert, daß für $n, m \in \mathbb{N}$ mit $n > m$ $u_r \rightarrow u_{T^m} \rightarrow u_{T^n}$ eine Markov-Kette ist. Das Data-processing-Lemma liefert:

$$I[u_r, u_{T^n}] \leq I[u_r, u_{T^m}]$$

$I[u_r, u_{T^n}]$ fällt monoton in n . Dies sichert die Existenz des Grenzwertes. Insbesondere ist wegen der Monotonie

$$\inf_{n \geq 1} I[u_r, u_{T^n}] = \lim_{n \rightarrow \infty} I[u_r, u_{T^n}]$$

Wir zeigen, daß obiger Grenzwert wesentlich von der Interaktionsstärke des Modells abhängt. Dazu verwenden wir die obere und untere Abschätzung aus den Sätzen 20 und 5.

Sei $\alpha > \alpha_{c,I} = \frac{\sqrt{\text{brT}}+1}{\sqrt{\text{brT}}-1}$. Damit ist

$$\theta^{-2} := \left(\frac{\alpha - 1}{\alpha + 1} \right)^{-2} < \left(\frac{\alpha_{c,I} - 1}{\alpha_{c,I} + 1} \right)^{-2} = \text{brT}$$

Nach Satz 1, der die Branching-Number eines Baumes mit einem elektrischen Netzwerk in Verbindung bringt, hat das Netzwerk auf T , in dem jede Kante die Leitfähigkeit

$$\bar{C}(e) := (\theta^{-2})^{-|e|} = \theta^{2|e|}$$

erhält, endlichen effektiven Widerstand. Vermöge des Rayleigh-Prinzips aus Satz 2 hat das Netzwerk, in dem jede Kante die Leitfähigkeit

$$C(e) := \frac{\theta^{2|e|}}{1 - \theta^2} > \theta^{2|e|} = \bar{C}(e)$$

erhält, ebenfalls endlichen effektiven Widerstand $R_{\text{eff}}(r, \partial T)$. Mit Satz 20 gilt:

$$\lim_{n \rightarrow \infty} I[u_r, u_{T^n}] \geq \lim_{n \rightarrow \infty} \frac{1}{2} \cdot \frac{1}{1 + R_{\text{eff}}(r, T^n)} > 0$$

Sei nun $\alpha < \alpha_{c,I}$. Damit ist $\theta = \frac{\alpha+1}{\alpha-1} > \text{br}T^{-1/2}$ und $\theta^{-2} < \text{br}T$. Mit der Definition der Branching-Number folgt:

$$\inf_{\Pi \subset T} \sum_{\sigma \in \Pi} \theta^{-2|\sigma|} = 0$$

Lemma 5 liefert:

$$\lim_{n \rightarrow \infty} I[u_r, u_{T^n}] = \inf_{n \geq 1} I[u_r, u_{T^n}] \leq \inf_{\Pi \in T} \sum_{\sigma \in \Pi} \theta^{-2|\sigma|} = 0 \quad (5.39)$$

□

5.2 Eine kritische Interaktionsstärke für Phasenübergang

Hier wollen wir uns etwas intensiver mit der Frage beschäftigen, wann die Randbedingungen der Plus-Phase den Spin der Wurzel beeinflussen.

Wir werden sehen, daß die beiden Fälle $\mathbf{P}^+[u_r = 1] = \mathbf{P}[u_r = 1]$ und $\mathbf{P}^+[u_r = 1] > \mathbf{P}[u_r = 1]$ durch einen kritischen Parameter getrennt sind.

Bevor wir mit dem Beweis beginnen, führen wir eine Funktion ein, die wir verwenden werden:

Lemma 6. *Sei $\alpha > 1$ und die Abbildung $g_\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}$ definiert als:*

$$g_\alpha : x \mapsto g_\alpha(x) := \frac{\alpha x + 1}{\alpha + x} \quad (5.40)$$

Weiterhin sei θ definiert als

$$\theta := \frac{\alpha - 1}{\alpha + 1} \quad (5.41)$$

Mit diesen Bezeichnungen gilt für alle $x \geq 1$

$$1 \leq g_\alpha(x) \leq \alpha \quad (5.42)$$

$$x^{\frac{\alpha-x}{\alpha+x}} \leq g_\alpha(x) \leq x^\theta \quad (5.43)$$

Abbildung 5.5 zeigt eine Skizze von g_α für $\alpha = 2$.

Beweis. Betrachten wir zuerst einmal die Ungleichung (5.42). Es gilt:

$$\begin{aligned} & x \geq 1 \\ \iff & x(1 - \alpha) \leq 1 - \alpha \\ \iff & x + \alpha \leq \alpha x + 1 \\ \iff & 1 \leq \frac{\alpha x + 1}{x + \alpha} = g_\alpha(x) \end{aligned} \quad (5.44)$$

Weiterhin ist $g_\alpha(x)$ für $x \geq 1$ streng monoton wachsend, denn es gilt:

$$g'_\alpha(x) = \frac{\alpha(x + \alpha) - (\alpha x + 1)}{(x + \alpha)^2} = \frac{\alpha^2 - 1}{(x + \alpha)^2} > 0 \quad (5.45)$$

Schließlich erhalten wir als Grenzwert für $x \rightarrow \infty$

$$\lim_{x \rightarrow \infty} g_\alpha(x) = \alpha \quad (5.46)$$

Insgesamt gilt also:

$$1 \leq g_\alpha(x) \leq \alpha$$

Wenden wir uns jetzt der Ungleichung (5.43) zu. Zunächst ist:

$$\begin{aligned} g_\alpha(x) &= \frac{\alpha x}{x + \alpha} + \frac{1}{x + \alpha} \\ &= \frac{\alpha}{x + \alpha} \cdot x + \frac{x}{x + \alpha} \cdot x^{-1} \end{aligned} \quad (5.47)$$

$$\geq x^{\frac{\alpha}{\alpha+x}} \cdot x^{-1 \cdot \frac{x}{\alpha+x}} \quad (5.48)$$

$$= x^{\frac{\alpha-x}{\alpha+x}}$$

Die Umformung von (5.47) nach (5.48) verwendet eine gewichtete Form der wohlbekannten Ungleichung zwischen geometrischem und arithmetischem Mittel, wie sie zum Beispiel in [4] zu finden ist.

Andererseits gilt für alle $x \geq 1$:

$$\begin{aligned} \log g_\alpha(x) &= \int_1^x \frac{d}{dt} (\log g_\alpha(t)) dt = \int_1^x \frac{1}{g_\alpha(t)} \cdot g'_\alpha(t) dt \\ &= \int_1^x \frac{t + \alpha}{\alpha t + 1} \cdot \frac{\alpha^2 - 1}{(t + \alpha)^2} dt = \int_1^x \frac{\alpha^2 - 1}{\alpha(t - 1)^2 + (\alpha + 1)^2 t} dt \\ &\leq \int_1^x \frac{\alpha^2 - 1}{(\alpha + 1)^2 t} dt = \int_1^x \frac{\alpha - 1}{\alpha + 1} \cdot \frac{1}{t} dt \\ &= \int_1^x \theta \cdot \frac{1}{t} dt = \log x^\theta \end{aligned}$$

Aufgrund der Monotonie der log-Funktion ist dies gleichbedeutend mit:

$$g_\alpha(x) \leq x^\theta$$

□

Wir werden nun die Aussage über die kritische Interaktionsstärke für Einfluß der Plus-Phase beweisen und dabei auch die Plus-Phase des Ising-Modells besser kennenlernen. Zunächst noch einmal die Aussage, die uns auf den nächsten Seiten beschäftigen wird.

Satz 16. *Sei $(\mathcal{U}^{K_T}, \mathcal{A}, \mathbf{P})$ das Ising-Modell mit Interaktionsstärke α und sei*

$$\alpha_{c,P} := \frac{\text{br}T + 1}{\text{br}T - 1}$$

Dann gilt

$$\mathbf{P}[u_r = 1] = \frac{1}{2} < \mathbf{P}^+[u_r = 1] \text{ , } \quad \text{falls } \alpha > \alpha_{c,P} \quad (5.49)$$

$$\mathbf{P}[u_r = 1] = \frac{1}{2} = \mathbf{P}^+[u_r = 1] \text{ , } \quad \text{falls } \alpha < \alpha_{c,P} \quad (5.50)$$

Beweis. Die Vorgabe der Randbedingungen in der Plus-Phase beeinflusst die Verteilung der η_e . Insbesondere geht die Unabhängigkeit der Zufallsvariablen verloren.

Wenn man eine Realisierung des Broadcasting-Modells mit der Randbedingung der Plus-Phase (auf einem endlichen Baum) konstruieren will, wird man mit den Kanten beginnen, die zum Rand führen und sich dann rekursiv zur Wurzel hinarbeiten.

Diese rekursive Vorgehensweise charakterisiert auch die folgende Beweisführung. Wir werden den Einfluß der Plus-Randbedingung auf einen Knoten σ als Funktion desselben Einflusses auf seine Nachfolger ausdrücken. Um die Aussage des Satzes zu beweisen, schätzen wir diesen Einfluß dann durch geeignete Flüsse auf T ab, und können so die Definition der Branching-Number ausnutzen.

Wir rekapitulieren kurz, wie wir in Definition 22 die Plus-Phase des Ising-Modells mit Temperatur t eingeführt haben.

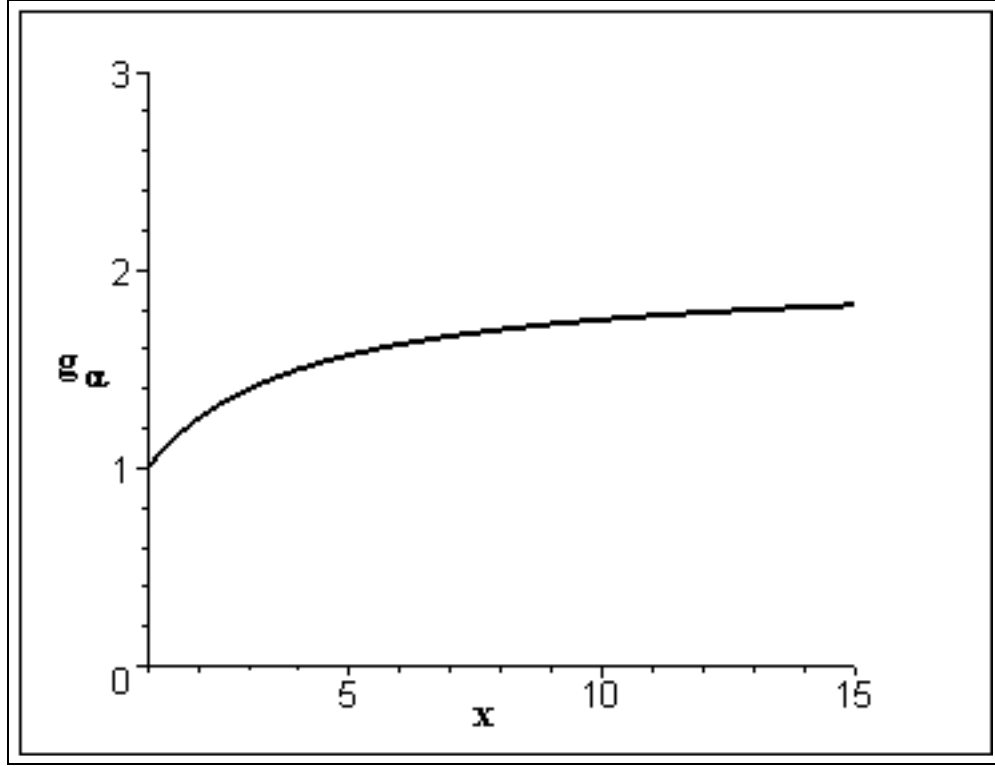


Abbildung 5.5: Der Graph von g_α für $\alpha = 2$

Die Menge aller möglichen Konfigurationen des Broadcasting-Modells auf $(\leq T^n)$ mit der Randbedingung $u_{T^n} \equiv +1$ ist:

$$\mathcal{U}^{n,+} := \{w \in \{-1, +1\}^{\leq T^n} : w|_{T^n} \equiv +1\}$$

Jeder Konfiguration aus $\mathcal{U}^{n,+}$ wird eine Energie $H(w)$ zugeordnet:

$$H(w) = -\frac{1}{t} \sum_{r < \sigma \leq T^n} w|_\sigma w|_{\bar{\sigma}}$$

Die Zustandssumme aller Konfigurationen $w \in \mathcal{U}^{n,+}$ ist:

$$Z^{n,+} := \sum_{w \in \mathcal{U}^{n,+}} \exp[-H[w]]$$

Mit diesen Bezeichnungen erhalten wir die folgende Darstellung für $\mathbf{P}^{n,+}$. Für

$A \subset \mathcal{U}^{n,+}$ ist:

$$\mathbf{P}^{n,+}[A] = \frac{\sum_{w \in A} \exp[-H[w]]}{Z^{n,+}}$$

Die Interaktionsstärke α zur Temperatur t ist wie in Definition 21:

$$\alpha = \exp\left[\frac{2}{t}\right] \quad (5.51)$$

Wir definieren nun eine Größe $Z_a^n(\sigma)$, die den Einfluß der +1-Randbedingung auf den Zustand a im Knoten σ mißt:

$$Z_a^n(\sigma) := \sum_{\substack{w \in \mathcal{U}^{n,+} \\ w_\sigma = a}} \exp \left[\frac{1}{t} \sum_{\sigma < \tau \leq S^n} w]_\tau w]_{\leftarrow \tau} \right]$$

Diese Größe wird uns helfen, eine Rekursionsgleichung aufzustellen, die den Einfluß des Randes auf einen Knoten σ durch den Einfluß auf die Nachfolger von σ ausdrückt. Es gilt nämlich:

$$\begin{aligned} Z_a^n(\sigma) &= \sum_{\substack{w \in \mathcal{U}^{n,+} \\ w_\sigma = a}} \exp \left[\frac{1}{t} \left(\sum_{\zeta \in \vec{\sigma}} \left(aw]_\zeta + \sum_{\zeta < \tau \leq S^n} w]_{\tau a u} w]_{\leftarrow \tau} \right) \right) \right] \\ &= \sum_{\substack{w \in \mathcal{U}^{n,+} \\ w_\sigma = a}} \left(\prod_{\zeta \in \vec{\sigma}} \exp \left[\frac{aw]_\zeta}{t} \right] \cdot \exp \left[\frac{1}{t} \sum_{\zeta < \tau \leq S^n} w]_\tau w]_{\leftarrow \tau} \right] \right) \\ &= \sum_{v \in \{-1, +1\}^{\vec{\sigma}}} \left(\prod_{\zeta \in \vec{\sigma}} \exp \left[\frac{av\zeta}{t} \right] \cdot Z_{v\zeta}^n(\zeta) \right) \\ &= \prod_{\zeta \in \vec{\sigma}} \left(\exp \left[\frac{a \cdot +1}{t} \right] Z_{+1}^n(\zeta) + \exp \left[\frac{a \cdot -1}{t} \right] Z_{-1}^n(\zeta) \right) \quad (5.52) \end{aligned}$$

Wir sind im besonderen daran interessiert, wie das Verhältnis der Randbedingung auf die Spins +1 respektive -1 im Knoten σ ist und definieren daher für alle Knoten $\sigma \in (< T^n)$:

$$f_n(\sigma) := \frac{Z_{+1}^n(\sigma)}{Z_{-1}^n(\sigma)}$$

Ist $f_n(\sigma) = 1$, so hat die Vorgabe der +1-Randbedingung auf der n -Sphäre keinen Einfluß auf den Knoten σ . f_n ist also ein geeignetes Werkzeug, um den Einfluß der Plus-Phase zu untersuchen. Mit Gleichung (5.52) und der Funktion g_α aus Lemma 6 können wir nun $f_n(\sigma)$ folgendermaßen darstellen:

$$\begin{aligned}
f_n(\sigma) &= \frac{Z_{+1}^n(\sigma)}{Z_{-1}^n(\sigma)} \\
&= \prod_{\zeta \in \overline{\sigma}} \frac{\exp[\frac{1}{t}] \cdot Z_{+1}^n(\zeta) + \exp[-\frac{1}{t}] \cdot Z_{-1}^n(\zeta)}{\exp[-\frac{1}{t}] \cdot Z_{+1}^n(\zeta) + \exp[\frac{1}{t}] \cdot Z_{-1}^n(\zeta)} \\
&= \prod_{\zeta \in \overline{\sigma}} g_\alpha \left(\frac{Z_{+1}^n(\zeta)}{Z_{-1}^n(\zeta)} \right) \\
&= \prod_{\zeta \in \overline{\sigma}} g_\alpha(f_n(\zeta)) \tag{5.53}
\end{aligned}$$

Für die Wurzel r des Baumes T gilt mit den obigen Bezeichnungen:

$$\begin{aligned}
f_n(r) &= \frac{Z_{+1}^n(r)}{Z_{-1}^n(r)} \\
&= \frac{Z_{+1}^n(r)/Z^{n,+}}{Z_{-1}^n(r)/Z^{n,+}} \\
&= \frac{\mathbf{P}^{n,+}[u_r = +1]}{\mathbf{P}^{n,+}[u_r = -1]}
\end{aligned}$$

Für $n \rightarrow \infty$, also im thermodynamischen Grenzübergang, erhalten wir:

$$f(r) := \lim_{n \rightarrow \infty} f_n(r) = \frac{\mathbf{P}^+[u_r = +1]}{\mathbf{P}^+[u_r = -1]} \tag{5.54}$$

Zum Beweis der Aussage aus Satz 15 ist die Betrachtung von f lohnenswert, denn es gilt:

$$\mathbf{P}^+[u_r = +1] = \frac{1}{2} = \mathbf{P}^+[u_r = -1] \quad \Longleftrightarrow \quad f(r) = 1$$

Wir werden bald sehen, daß das Verhalten von f wesentlich von der Interaktionsstärke des Modells abhängt, denn es gilt:

$$f(r) = 1 \quad \Longleftrightarrow \quad \alpha < \alpha_{c,P} \tag{5.55}$$

Damit ist dann auch Satz 15 bewiesen.

Um die gesuchte Eigenschaft von $f(r)$ nachzuweisen, werden wir f_n mittels geeigneter Flüsse μ_Π und ϕ_Π auf T abschätzen. Wir werden dabei so vorgehen, daß wir für einen geeigneten Cutset Π die Werte der Flüsse so vorgeben, daß die gesuchte Abschätzung gilt.

Sei zunächst $\alpha < \alpha_{c,P} = \frac{\text{br}T+1}{\text{br}T-1}$.

Mit der Definition von θ aus (5.41) und der Definition der Branching-Number aus 1 folgt:

$$\alpha < \alpha_{c,P} \implies \theta < \frac{1}{\text{br}T} \implies \inf_{\Pi \subset K_T} \sum_{\sigma \in \Pi} \theta^{|\sigma|} = 0 \quad (5.56)$$

Sei $n \in \mathbb{N}$ fest gewählt und sei $\Pi \subset (< T^n)$ ein Cutset in T . Wir definieren nun $h_\Pi : (\leq \Pi) \rightarrow \mathbb{R}_0^+$ wie folgt

$$h_\Pi(\sigma) = \begin{cases} \alpha^{\frac{1}{\theta}} & \text{für } \sigma \in \Pi \\ \prod_{\tau \in \bar{\sigma}} h_\Pi(\tau)^\theta & \text{für } \sigma \in (< \Pi) \end{cases} \quad (5.57)$$

Für $\sigma \in \Pi$ gilt unter Verwendung von Ungleichung (5.42):

$$g_\alpha(f_n(\sigma)) \leq \alpha = h_\Pi(\sigma)^\theta \quad (5.58)$$

Für $\sigma \in (< \Pi)$ gilt weiterhin mit Ungleichung (5.43):

$$g_\alpha(f_n(\sigma)) \leq f_n(\sigma)^\theta \quad (5.59)$$

h_Π ist somit auf den Knoten, die zwischen r und Π liegen, eine obere Abschätzung für f_n . Dies folgt aus den letzten beiden Ungleichungen, wenn man die Darstellung von f_n aus (5.53) benutzt. Damit gilt auch:

$$f_n(r) \leq h_\Pi(r) \quad (5.60)$$

Wir benutzen h_Π , um einen Fluß μ_Π von r nach Π zu definieren:

$$\mu_\Pi(\sigma) := \theta^{|\sigma|} \log h_\Pi(\sigma)$$

μ_Π ist ein Fluß auf $(< \Pi)$, denn für $\sigma \in (< \Pi)$ gilt:

$$\begin{aligned}
\mu_\Pi(\sigma) &= \theta^{|\sigma|} \log \left[\prod_{\tau \in \vec{\sigma}} h_\Pi(\tau)^\theta \right] \\
&= \theta^{|\sigma|} \cdot \left(\sum_{\tau \in \vec{\sigma}} \log [h_\Pi(\tau)^\theta] \right) \\
&= \theta^{|\sigma|} \cdot \theta \cdot \sum_{\tau \in \vec{\sigma}} \log [h_\Pi(\tau)] \\
&= \sum_{\tau \in \vec{\sigma}} \mu_\Pi(\tau)
\end{aligned}$$

Wegen der Baumstruktur von T genügt dies, um sicherzustellen, daß μ_Π ein Fluß von r nach Π ist. Wir verwenden diese Eigenschaft von μ_Π und erhalten:

$$\begin{aligned}
h_\Pi(r) &= \exp [\mu_\Pi(r)] \\
&= \exp \left[\sum_{\tau \in \Pi} \mu_\Pi(\tau) \right] \\
&= \exp \left[\sum_{\tau \in \Pi} \theta^{|\tau|} \log [\alpha^{\frac{1}{\theta}}] \right] \\
&= \alpha^{\left(\frac{1}{\theta} \sum_{\tau \in \Pi} \theta^{|\tau|} \right)}
\end{aligned}$$

Wie in (5.56) bereits bemerkt, ist θ für $\alpha < \alpha_{c,P}$ kleiner als $\text{br}T^{-1}$ und mit der Definition der Branching-Number folgt damit:

$$\inf_{\Pi \subset K_T} h_\Pi(r) = \inf_{\Pi \subset K_T} \alpha^{\left(\frac{1}{\theta} \sum_{\tau \in \Pi} \theta^{|\tau|} \right)} = \alpha^0 = 1 \quad (5.61)$$

Gleichung (5.60) gilt für alle Cutsets $\Pi \subset (< T^n)$. Für $n \rightarrow \infty$ erhalten wir somit:

$$1 \leq f(r) = \lim_{n \rightarrow \infty} f_n(r) \leq \inf_{\Pi \subset K_T} h_\Pi(r) \leq 1 \quad (5.62)$$

Insgesamt haben wir also gezeigt $\alpha < \alpha_{c,P} \implies f(r) = 1$.

Nun wenden wir uns der Rückrichtung der Aussage zu. Sei $f(r) = 1$ und $\alpha \in \mathbb{R}^+$ die Interaktionsstärke des Modells. Wir werden zeigen, daß $\alpha < \alpha_{c,P}$ ist.

Sei $\alpha_0 \in (0, \alpha]$ und wie in (5.41) sei

$$\theta_0 := \frac{\alpha_0 - 1}{\alpha_0 + 1} \iff \alpha_0 = \frac{1 + \theta_0}{1 - \theta_0} \quad (5.63)$$

Wir werden zeigen, daß $\theta_0 < \text{brT}^{-1}$ ist. Hieraus folgt mit (5.63) unmittelbar:

$$\alpha_0 < \frac{\text{brT} + 1}{\text{brT} - 1} = \alpha_{c,P}$$

Wir gehen bei diesem Nachweis ähnlich vor wie bei der schon gezeigten Richtung der Äquivalenzaussage. Wir werden einen Fluß $\phi_{\tilde{\Pi}}$ definieren, dessen Werte wir für einen geeigneten Cutset $\tilde{\Pi}$ vorgeben.

Wir befinden uns in der Plus-Phase des attraktiven Ising-Modells. Daher ist für alle $\sigma \in \mathcal{T}$ $f(\sigma) \geq 1$. Die Voraussetzung in dieser Beweisrichtung der Äquivalenz ist $f(r) = 1$. Mit der Rekursionsgleichung (5.53) folgt, daß für alle $\sigma \in \mathcal{T}$ gilt:

$$f(\sigma) = 1 \quad (5.64)$$

Wir hatten $\alpha_0 \in (0, \alpha]$ gewählt. Damit ist $\frac{\alpha}{\alpha_0} \geq 1$. Da nun

$$\lim_{n \rightarrow \infty} f_n(\sigma) = f(\sigma) = 1$$

ist, folgt, daß es ein $n_0 \in \mathbb{N}$ gibt, so daß für alle $n > n_0$ und für alle $\sigma \in \mathcal{T}^1$ gilt:

$$f_n(\sigma) \leq \frac{\alpha}{\alpha_0}$$

Im weiteren sei nun $n > n_0$ fest gewählt. Damit ist es möglich, einen speziellen Cutset $\tilde{\Pi}$ zu wählen, der folgende Eigenschaften erfüllt:

$$\begin{aligned} \forall \sigma \in (\leq \tilde{\Pi}) : & \quad f_n(\sigma) \leq \frac{\alpha}{\alpha_0} \\ \forall \sigma \in \tilde{\Pi} \quad \exists \tau_0 \in \vec{\sigma} : & \quad f_n(\tau_0) > \frac{\alpha}{\alpha_0} \end{aligned}$$

Nach Konstruktion ist $\tilde{\Pi} \subset (\leq T^n)$.

Um eine untere Abschätzung $h_{\tilde{\Pi}}$ für f_n angeben zu können, setzen wir

$$a := g_\alpha \left(\frac{\alpha}{\alpha_0} \right)$$

Mit dieser Bezeichnung sei $h_{\tilde{\Pi}} : (\leq \tilde{\Pi}) \rightarrow \mathbb{R}^+$ definiert als:

$$h_{\tilde{\Pi}}(\sigma) := \begin{cases} a & \text{falls } \sigma \in \tilde{\Pi} \\ \prod_{\tau \in \vec{\sigma}} h_{\tilde{\Pi}}(\tau)^{\theta_0} & \text{falls } \sigma \in < (\tilde{\Pi}) \end{cases} \quad (5.65)$$

Um zu zeigen, daß $h_{\tilde{\Pi}}$ eine untere Abschätzung für f_n ist, betrachten wir die Fälle $\sigma \in \Pi$ und $\sigma \in (< \Pi)$ getrennt.

Sei $\sigma \in \tilde{\Pi}$. Wir befinden uns in der Plus-Phase des Ising-Modells, daher gilt $f_n(\tau) \geq 1$ für alle $\tau \in T$. Mit den Eigenschaften von g_α aus Lemma 6 folgt:

$$f_n(\sigma) = \prod_{\tau \in \vec{\sigma}} g_\alpha(f_n(\tau)) \geq g_\alpha(f_n(\tau_0)) > a = h_{\tilde{\Pi}}(\sigma) \quad (5.66)$$

Sei nun $\sigma \in (< \tilde{\Pi})$. Damit ist $x := f_n(\sigma) > \frac{\alpha}{\alpha_0}$ und mit (5.43) folgt:

$$\begin{aligned} g_\alpha(f_n(\sigma)) &= g_\alpha(x) \\ &\leq x^\theta \\ &= f_n(\sigma)^\theta \end{aligned} \quad (5.67)$$

Die Kombination aus (5.66) und (5.67) zeigt, daß für alle $\sigma \in (\leq \tilde{\Pi})$ gilt:

$$h_{\tilde{\Pi}}(\sigma) \leq f_n(\sigma)$$

$h_{\tilde{\Pi}}$ ist eine untere Abschätzung für f_n . Nun benutzen wir $h_{\tilde{\Pi}}$, um einen Fluß $\phi_{\tilde{\Pi}}$ von r nach $\tilde{\Pi}$ zu definieren:

$$\phi_{\tilde{\Pi}}(\sigma) := \theta_0^{|\sigma|} \log[h_{\tilde{\Pi}}(\sigma)]$$

$\phi_{\tilde{\Pi}}$ ist ein Fluß, denn es gilt:

$$\begin{aligned}
\phi_{\tilde{\Pi}} &= \theta_0^{|\sigma|} \log [h_{\tilde{\Pi}}(\sigma)] \\
&= \theta_0^{|\sigma|} \log \left[\prod_{\tau \in \vec{\sigma}} h_{\tilde{\Pi}}(\tau)^{\theta_0} \right] \\
&= \theta_0^{|\sigma|} \sum_{\tau \in \vec{\sigma}} \log [h_{\tilde{\Pi}}(\tau)^{\theta_0}] \\
&= \theta_0^{|\sigma|+1} \sum_{\tau \in \vec{\sigma}} \log [h_{\tilde{\Pi}}(\tau)] \\
&= \sum_{\tau \in \vec{\sigma}} \theta_0^{|\tau|} \log [h_{\tilde{\Pi}}(\tau)] \\
&= \sum_{\tau \in \vec{\sigma}} \phi_{\tilde{\Pi}}(\tau)
\end{aligned}$$

Wir können somit $h_{\tilde{\Pi}}(r)$ wie folgt schreiben:

$$\begin{aligned}
h_{\tilde{\Pi}}(r) &= \exp [\phi_{\tilde{\Pi}}(r)] \\
&= \exp \left[\sum_{\tau \in \tilde{\Pi}} \phi_{\tilde{\Pi}}(\tau) \right] \\
&= \exp \left[\sum_{\tau \in \tilde{\Pi}} \theta_0^{|\tau|} \log(a) \right] \\
&= a^{\left(\sum_{\tau \in \tilde{\Pi}} \theta_0^{|\tau|} \right)}
\end{aligned}$$

Bei diesen Umformungen haben wir benutzt, daß die Stärke eines Flusses gleich der Summe des Flusses durch jeden Cutset ist. Für unsere spezielle Wahl des Cutsets $\tilde{\Pi}$ gilt somit:

$$1 \leq h_{\tilde{\Pi}}(r) = a^{\left(\sum_{\tau \in \tilde{\Pi}} \theta_0^{|\tau|} \right)} \leq f_n(r)$$

Für $n \rightarrow \infty$ wird dies zu

$$1 \leq a \left(\sum_{\tau \in \tilde{\Pi}} \theta_0^{|\tau|} \right) \leq f(r) = 1$$

Da $a > 1$ ist, bedeutet dies, daß

$$\sum_{\tau \in \tilde{\Pi}} \theta_0^{|\tau|} = 0$$

ist. Insbesondere verschwindet also das Infimum über alle Cutsets $\Pi \subset T$:

$$\inf_{\Pi \subset T} \sum_{\tau \in \Pi} \theta_0^{|\tau|} = \inf_{\Pi \subset T} \sum_{\tau \in \Pi} \left(\frac{1}{\theta_0} \right)^{-|\tau|} = 0$$

Nach Definition der Branching-Number ist nun sofort klar, daß $\frac{1}{\theta_0} > \text{br}T$ ist, womit auch $\alpha_0 < \alpha_{c,P}$ ist.

Wir haben gezeigt

$$f(r) = 1 \iff \alpha < \alpha_{c,P}$$

Nach Konstruktion von f ist damit die Aussage bewiesen. □

Kapitel 6

Anhang

6.1 Lyons Nachweis der kritischen Interaktion für Spinperkolation in der Plus-Phase

Wir geben hier einen zweiten Beweis von Satz 19 an, der die kritische Interaktionsstärke für Spinperkolation in der Plus-Phase des Ising-Modells benennt. Dieser Beweis benutzt keine Aussagen über die kritischen Parameter im freien Ising-Modell. Er entspricht dem Beweis, den Lyons in [7] führt, um eine kritische Temperatur für Spinperkolation in der Plus-Phase zu berechnen.

Im Fall freier Randbedingungen war es sehr einfach, die Aussage über den kritischen Parameter für Spinperkolation zu beweisen. Dies beruhte auf der Tatsache, daß in diesem Fall die Spinperkolation eine Bernoulli-Perkolation ist. In der Plus-Phase ist die Spinperkolation allerdings keine Bernoulli-Perkolation mehr, die Vorgabe von Randbedingungen zerstört die Unabhängigkeit der η_e . Wir führen eine größere Klasse von Perkolationen ein, in der wir auch die Spinperkolation in der Plus-Phase wiederfinden werden.

6.1.1 Quasi-Bernoulli-Perkolationen

Definition 26 ([7]). $(\mathcal{T}, \mathbf{P})$ heißt **quasi-Bernoulli-Perkolation**, wenn es eine positive Zahl M gibt, so daß für alle $\sigma, \tau \in K_{\mathcal{T}}$ gilt:

$$\begin{aligned} \mathbf{P}[(\sigma \leftrightarrow r), (\tau \leftrightarrow r) \mid (\sigma \wedge \tau \leftrightarrow r)] &\leq \\ &\leq M \cdot \mathbf{P}[(\sigma \leftrightarrow r) \mid (\sigma \wedge \tau \leftrightarrow r)] \cdot \mathbf{P}[(\tau \leftrightarrow r) \mid (\sigma \wedge \tau \leftrightarrow r)] \end{aligned} \quad (6.1)$$

Die Bernoulli-Perkolationen sind natürlich mit Gleichheit und $M = 1$ in der Klasse der quasi-Bernoulli-Perkolationen enthalten.

Durch einfache Umformungen erhalten wir aus (6.1) zwei äquivalente Aussagen, die wir später benutzen werden. Mit den gleichen Vorgaben wie oben gilt für eine quasi-Bernoulli-Perkolation für alle $\sigma, \tau \in K_T$:

$$\begin{aligned} \mathbf{P}[(\sigma \leftrightarrow r) \cap (\tau \leftrightarrow r)] \cdot \mathbf{P}[\sigma \wedge \tau \leftrightarrow r] &\leq M \mathbf{P}[\sigma \leftrightarrow r] \mathbf{P}[\tau \leftrightarrow r] \\ \iff \mathbf{P}[\sigma \leftrightarrow r]^{-1} \mathbf{P}[\tau \leftrightarrow r]^{-1} \mathbf{P}[(\sigma \leftrightarrow r) \cap (\tau \leftrightarrow r)] &\leq M \cdot \mathbf{P}[\sigma \wedge \tau \leftrightarrow r]^{-1} \end{aligned}$$

Das folgende Lemma macht eine Aussage darüber, wann eine unendlichgroße Zusammenhangskomponente der Wurzel in einer quasi-Bernoulli-Perkolation positive Wahrscheinlichkeit hat:

Lemma 7 ([8], Theorem 3.1). *Sei $(\mathcal{T}, \mathbf{P})$ eine quasi-Bernoulli-Perkolation. Falls*

$$\inf_{\Pi} \sum_{\sigma \in \Pi} \mathbf{P}[\sigma \leftrightarrow r] = 0 \quad (6.2)$$

ist, dann ist $\mathbf{P}[|\hat{T}_r| < \infty] = 1$.

Gibt es hingegen positive Zahlen w_n mit $\sum_{n \geq 1} w_n < \infty$, und gilt weiterhin

$$\inf_{\Pi} \sum_{\sigma \in \Pi} w_{|\sigma|} \mathbf{P}[\sigma \leftrightarrow r] > 0 \quad (6.3)$$

so ist $\mathbf{P}[|\hat{T}_r| = \infty] > 0$.

Während der erste Teil der Aussage leicht einsichtig ist und der Beweis nur aus wenigen technischen Details besteht, ist die zweite Aussage nicht sofort klar. Ihr Beweis benutzt eine Aussage aus der Theorie der elektrischen Netzwerke auf Graphen und folgende Aussagen über uniform integrierbare Folgen von Zufallsvariablen:

Satz 26 ([12], 13.3a). *Eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsvariablen ist uniform integrierbar, falls es eine Zahl $K \in [0, \infty)$ und ein $s > 1$ gibt, so daß für alle $n \in \mathbb{N}$ gilt:*

$$\mathbf{E}[|X_n|^s] < K \quad (6.4)$$

Satz 27 ([12], 13.7). *Sei X eine Zufallsvariable und $(X_n)_{n \in \mathbb{N}}$ eine zufällige Folge. Dann gilt*

$$\mathbf{E}[|X_n - X|] \xrightarrow{n \rightarrow \infty} 0 \quad (6.5)$$

genau dann, wenn beide folgenden Aussagen erfüllt sind:

$$a) \quad X_n \rightarrow X \quad \text{in Wahrscheinlichkeit} \quad (6.6)$$

$$b) \quad X_n \quad \text{ist uniform integrierbar} \quad (6.7)$$

Beweis von Lemma 7. Aus (6.2) folgt, daß es eine Folge $\{\Pi_n\}$ von Cutsets gibt, für die gilt:

$$\begin{aligned} \sum_{\sigma \in \Pi_n} \mathbf{P}[\sigma \leftrightarrow r] &\xrightarrow{n \rightarrow \infty} 0 \\ \iff \mathbf{E}[|(\hat{T}_r \cap \Pi_n)|] &\xrightarrow{n \rightarrow \infty} 0 \end{aligned} \quad (6.8)$$

Mit dem Lemma von Fatou folgt aus (6.8):

$$\begin{aligned} \liminf_{n \rightarrow \infty} |(\hat{T}_r \cap \Pi_n)| &= 0 && \mathbf{P}\text{-fast sicher} \\ \implies |\hat{T}_r| &< \infty && \mathbf{P}\text{-fast sicher} \end{aligned} \quad (6.9)$$

Um die andere Aussage des Satzes zu beweisen, betrachten wir das elektrische Netzwerk auf T , in dem jeder Kante die Leitfähigkeit

$$C(e_{(\bar{\sigma}, \sigma)}) = \mathbf{P}[\sigma \leftrightarrow r]$$

zugeordnet wird.

Sei μ nun ein Einheitsfluß von r zur Krone von T , der den durch die $w_{|e|}$ gemachten Vorgaben $\mu(e) \leq w_{|e|}$ genügt. Nach Satz 4 hat μ endliche Energie:

$$E(\mu) = \sum_{e \in T} \mu(e)^2 R(e) = \sum_{\sigma \in T} \mu(\sigma)^2 \mathbf{P}[\sigma \leftrightarrow r]^{-1} < \infty. \quad (6.10)$$

Wir konstruieren nun eine Folge X_n von Zufallsvariablen, mit deren Hilfe wir die gesuchte Aussage zeigen können:

$$X_n := \sum_{\sigma \in T^n} \mathbf{1}_{\{\sigma \leftrightarrow r\}} \mu(\sigma) \mathbf{P}[\sigma \leftrightarrow r]^{-1} \quad (6.11)$$

Die Zufallsvariablen X_n haben für alle $n \in \mathbb{N}$ Erwartungswert 1:

$$\begin{aligned}
\mathbf{E}[X_n] &= \mathbf{E} \left[\sum_{\sigma \in T^n} \mathbf{1}_{\{\sigma \leftrightarrow r\}} \mu(\sigma) \mathbf{P}[\sigma \leftrightarrow r]^{-1} \right] \\
&= \sum_{\sigma \in T^n} \mathbf{E} [\mathbf{1}_{\{\sigma \leftrightarrow r\}} \mu(\sigma) \mathbf{P}[\sigma \leftrightarrow r]^{-1}] \\
&= \sum_{\sigma \in T^n} \mathbf{P}[\sigma \leftrightarrow r] \mu(\sigma) \mathbf{P}[\sigma \leftrightarrow r]^{-1} \\
&= \sum_{\sigma \in T^n} \mu(\sigma) \\
&= 1
\end{aligned} \tag{6.12}$$

Weiterhin hat X_n auch ein beschränktes zweites Moment:

$$\begin{aligned}
\mathbf{E}[X_n^2] &= \mathbf{E} \left[\left(\sum_{\sigma \in T^n} \mathbf{1}_{\{\sigma \leftrightarrow r\}} \mu(\sigma) \mathbf{P}[\sigma \leftrightarrow r]^{-1} \right)^2 \right] \\
&= \mathbf{E} \left[\sum_{\sigma, \tau \in T^n} \mathbf{1}_{\{(\sigma \leftrightarrow r) \cap (\tau \leftrightarrow r)\}} \mu(\sigma) \mu(\tau) \mathbf{P}[\sigma \leftrightarrow r]^{-1} \mathbf{P}[\tau \leftrightarrow r]^{-1} \right] \\
&= \sum_{\sigma, \tau \in T^n} \mu(\sigma) \mu(\tau) \mathbf{P}[\sigma \leftrightarrow r]^{-1} \mathbf{P}[\tau \leftrightarrow r]^{-1} \mathbf{P}[(\sigma \leftrightarrow r) \cap (\tau \leftrightarrow r)]
\end{aligned} \tag{6.13}$$

$$\leq M \sum_{\sigma, \tau \in T^n} \mu(\sigma) \mu(\tau) \mathbf{P}[\sigma \wedge \tau \leftrightarrow r]^{-1} \tag{6.14}$$

$$\begin{aligned}
&= M \sum_{|\psi| \leq n} (\mathbf{P}[\psi \leftrightarrow r]^{-1} \sum_{\sigma, \tau \in T^n, \psi = \sigma \wedge \tau} \mu(\sigma) \mu(\tau)) \\
&\leq M \sum_{|\psi| \leq n} \mathbf{P}[\psi \leftrightarrow r]^{-1} \mu(\psi) \mu(\psi)
\end{aligned} \tag{6.15}$$

$$< \infty$$

Die Umformung von Gleichung (6.13) nach Gleichung (6.14) benutzt die Tatsache, daß es sich um eine quasi-Bernoulli-Perkolation handelt. Die darauf

folgende Umformung ist möglich, da μ ein Fluß auf T ist.

Da die Folge X_n ein beschränktes zweites Moment hat, ist sie gleichmäßig integrierbar.

Wir zeigen, daß eine unendlichgroße Zusammenhangskomponente der Wurzel positive Wahrscheinlichkeit hat. Dies geschieht mit einem Widerspruchsargument.

Annahme: $|\hat{T}_r| < \infty$ \mathbf{P} -fast sicher.

Dann ist auch $X_n \rightarrow 0$ \mathbf{P} -fast sicher. Aus der uniformen Integrierbarkeit der X_n folgt nun mit Satz 27 $\mathbf{E}[X_n] \xrightarrow{n \rightarrow \infty} 0$. Dies steht im Widerspruch zu Gleichung (6.12), die besagt, daß der Erwartungswert aller X_n 1 ist.

Die Annahme war falsch, also ist:

$$\mathbf{P}[|\hat{T}_r| = \infty] > 0 \quad (6.16)$$

□

Dieses Lemma gibt uns ein Kriterium an die Hand, um quasi-Bernoulli-Perkolationen nach unendlichgroßen Clustern zu untersuchen.

Im folgenden werden wir zeigen, daß Spinperkolation in der Plus-Phase des Broadcasting-Modells eine quasi-Bernoulli-Perkolation darstellt, und mit Hilfe von Lemma 7 die Aussage über die Existenz von unendlichgroßen Zusammenhangskomponenten der Wurzel nachweisen.

6.1.2 Kritische Interaktion für Spinperkolation

Wir beweisen die Aussage über die kritische Interaktionsstärke für Spinperkolation in der Plus-Phase des Broadcasting-Modells, die in folgendem Satz gemacht wird:

Satz 19. *Sei $(\mathcal{T}, \mathbf{P}^+)$ die Spinperkolation in der Plus-Phase des Broadcasting-Modells mit Interaktionsstärke α . Dann gilt*

$$\mathbf{P}^+[|\hat{T}_r| = \infty] = 0, \quad \text{falls } \alpha < \alpha_{c,Perk.}$$

$$\mathbf{P}^+[|\hat{T}_r| = \infty] = 1, \quad \text{falls } \alpha > \alpha_{c,Perk.}$$

Dabei ist

$$\alpha_{c,Perk.} = \max \left\{ \frac{1}{\text{br}T - 1}, 1 \right\}$$

Beweis. Das Broadcasting-Modell ist, wie wir gesehen haben, ein Markov-Feld, also üben nur direkte Nachbarn Einfluß aufeinander aus. Da T keine Zyklen besitzt, gibt es genau einen Pfad, der zwei Knoten σ und τ miteinander verbindet.

Beide Eigenschaften zusammen bedeuten, daß sich im Broadcasting-Modell die Spins zweier Knoten allenfalls über den Spin des jüngsten gemeinsamen Vorfahrens $\sigma \wedge \tau$ beeinflussen können.

Bedingt auf den Spin in diesem jüngsten gemeinsamen Vorfahren sind die Ereignisse auf dem Pfad zwischen $\sigma \wedge \tau$ und σ unabhängig von denen auf dem Pfad zwischen $\sigma \wedge \tau$ und τ . Diese Eigenschaft hängt nicht von etwaigen Randbedingungen ab, gilt also auch in der Plus-Phase. Daher ist:

$$\mathbf{P}^+[(\sigma \leftrightarrow \sigma \wedge \tau) \cap (\tau \leftrightarrow \sigma \wedge \tau)] = \mathbf{P}^+[\sigma \leftrightarrow \sigma \wedge \tau] \cdot \mathbf{P}^+[\tau \leftrightarrow \sigma \wedge \tau] \quad (6.17)$$

Dies gewährleistet, daß die Spinperkolation in der Plus-Phase eine quasi-Bernoulli-Perkolation ist, denn es gilt für alle $\sigma, \tau \in K_T$:

$$\begin{aligned} \mathbf{P}^+[(\sigma \leftrightarrow r) \cap (\tau \leftrightarrow r) | \sigma \wedge \tau \leftrightarrow r] \\ &= \mathbf{P}^+[(\sigma \leftrightarrow \sigma \wedge \tau) \cap (\tau \leftrightarrow \sigma \wedge \tau)] \\ &= \mathbf{P}^+[\sigma \leftrightarrow \sigma \wedge \tau] \cdot \mathbf{P}^+[\tau \leftrightarrow \sigma \wedge \tau] \\ &= \mathbf{P}^+[\sigma \leftrightarrow r | \sigma \wedge \tau \leftrightarrow r] \mathbf{P}^+[\tau \leftrightarrow r | \sigma \wedge \tau \leftrightarrow r] \end{aligned} \quad (6.18)$$

Wir können also Lemma 7 verwenden, um die Spinperkolation in der Plus-Phase zu untersuchen.

Wir leiten nun zuerst einen Ausdruck für $\mathbf{P}^+[\sigma \leftrightarrow r]$ her:

$$\begin{aligned} \mathbf{P}^+[\sigma \leftrightarrow r] &= \mathbf{P}^+[u_r = +1] \cdot \prod_{\tau \in \text{Pfad}(\sigma)} \mathbf{P}^+[u_\tau = +1 | u_{\tau^-} = +1] \\ &\quad + \mathbf{P}^+[u_r = -1] \cdot \prod_{\tau \in \text{Pfad}(\sigma)} \mathbf{P}^+[u_\tau = -1 | u_{\tau^-} = -1] \end{aligned} \quad (6.19)$$

Die nächsten Überlegungen werden das Ziel haben, einen Ausdruck für $\mathbf{P}^+[u_\tau = +1 | u_{\tau^-} = +1]$ und $\mathbf{P}^+[u_\tau = -1 | u_{\tau^-} = -1]$ abzuleiten. Wir definieren:

$$f(\tau) := \frac{\mathbf{P}^+[u_\tau = +1]}{\mathbf{P}^+[u_\tau = -1]} \quad (6.20)$$

Mit diesen Bezeichnungen und $\alpha = \frac{1-\epsilon}{\epsilon}$ erhalten wir:

$$\begin{aligned}
\frac{\mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = +1]}{\mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = -1]} &= \frac{\mathbf{P}^+[u_{\tau} = +1] \cdot \mathbf{P}^+[u_{\overline{\tau}} = +1 \mid u_{\tau} = +1]}{\mathbf{P}^+[u_{\tau} = -1] \cdot \mathbf{P}^+[u_{\overline{\tau}} = +1 \mid u_{\tau} = -1]} \\
&= f(\tau) \frac{\mathbf{P}^+[u_{\overline{\tau}} = +1 \mid u_{\tau} = +1]}{\mathbf{P}^+[u_{\overline{\tau}} = +1 \mid u_{\tau} = -1]} \\
&= f(\tau) \cdot \left(\frac{1-\epsilon}{\epsilon} \right) \\
&= f(\tau) \cdot \alpha
\end{aligned} \tag{6.21}$$

Auf gleiche Art und Weise berechnen wir:

$$\begin{aligned}
\frac{\mathbf{P}^+[u_{\overline{\tau}} = -1 \cap u_{\tau} = -1]}{\mathbf{P}^+[u_{\overline{\tau}} = -1 \cap u_{\tau} = +1]} &= \frac{\mathbf{P}^+[u_{\tau} = -1] \cdot \mathbf{P}^+[u_{\overline{\tau}} = -1 \mid u_{\tau} = -1]}{\mathbf{P}^+[u_{\tau} = +1] \cdot \mathbf{P}^+[u_{\overline{\tau}} = -1 \mid u_{\tau} = +1]} \\
&= \frac{1}{f(\tau)} \cdot \alpha
\end{aligned} \tag{6.22}$$

Unter Verwendung von (6.21) erhalten wir nun einen der gesuchten Ausdrücke:

$$\begin{aligned}
&\mathbf{P}^+[u_{\tau} = +1 \mid u_{\overline{\tau}} = +1] \\
&= \frac{\mathbf{P}^+[u_{\tau} = +1 \cap u_{\overline{\tau}} = +1]}{\mathbf{P}^+[u_{\overline{\tau}} = +1]} \\
&= \frac{\mathbf{P}^+[u_{\tau} = +1 \cap u_{\overline{\tau}} = +1]}{\mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = +1] + \mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = -1]} \\
&= \frac{\mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = +1]}{\mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = -1]} \\
&\quad \cdot \left(\frac{\mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = +1] + \mathbf{P}^+[u_{\overline{\tau}} = +1 \cap u_{\tau} = -1]}{\mathbf{P}^+[u_{\tau} = 1 \cap u_{\tau} = -1]} \right)^{-1} \\
&= \frac{f(\tau) \cdot \alpha}{f(\tau) \cdot \alpha + 1}
\end{aligned} \tag{6.23}$$

Mit einer analogen Rechnung unter Verwendung von (6.22) gilt:

$$\mathbf{P}^+[u_{\tau} = -1 \mid u_{\overline{\tau}} = -1] = \frac{f(\tau)^{-1} \cdot \alpha}{f(\tau)^{-1} \cdot \alpha + 1} \tag{6.24}$$

Unter Verwendung dieser Ergebnisse wird (6.19) zu

$$\begin{aligned} \mathbf{P}^+[\sigma \leftrightarrow r] &= \mathbf{P}^+[u_r = +1] \prod_{\tau \in \text{Pfad}(\sigma)} \frac{f(\tau) \cdot \alpha}{f(\tau) \cdot \alpha + 1} \\ &\quad + \mathbf{P}^+[u_r = -1] \prod_{\tau \in \text{Pfad}(\sigma)} \frac{f(\tau)^{-1} \cdot \alpha}{f(\tau)^{-1} \cdot \alpha + 1} \end{aligned} \quad (6.25)$$

Betrachten wir nun den Fall $\alpha < \alpha_{c,Perk.}$.

Da im Broadcasting-Modell die Interaktion stets größer oder gleich 1 ist, bedeutet $\alpha < \alpha_{c,Perk.}$, daß wir uns hier nur mit Bäumen mit $\text{brT} < 2$ beschäftigen, und somit $\alpha < \frac{1}{\text{brT}-1}$ ist. Dies bedeutet:

$$\frac{\alpha}{\alpha + 1} < \frac{\alpha_{c,Perk.}}{\alpha_{c,Perk.} + 1} = \frac{1}{\text{brT}} \quad (6.26)$$

Aus dem Vergleich der kritischen Interaktionsstärken aus Satz 18 folgt, daß $\alpha < \alpha_{c,P}$ ist. Für solche Interaktionen hat die Randbedingung der Plus-Phase keinen Einfluß auf das Innere des Baumes, es gilt:

$$\mathbf{P}^+[u_\tau = +1] = \mathbf{P}^+[u_\tau = -1] = \frac{1}{2} \quad (6.27)$$

Demnach ist $f(\tau) = 1$ für alle Knoten τ . Damit vereinfacht sich Gleichung (6.25) zu

$$\begin{aligned} \mathbf{P}^+[\sigma \leftrightarrow r] &= \mathbf{P}^+[u_r = 1] \prod_{\tau \in \text{Pfad}(\sigma)} \frac{f(\tau) \cdot \alpha}{f(\tau) \cdot \alpha + 1} \\ &\quad + \mathbf{P}^+[u_r = -1] \prod_{\tau \in \text{Pfad}(\sigma)} \frac{f(\tau)^{-1} \cdot \alpha}{f(\tau)^{-1} \cdot \alpha + 1} \\ &= \frac{1}{2} \cdot \prod_{\tau \in \text{Pfad}(\sigma)} \frac{\alpha}{\alpha + 1} + \frac{1}{2} \cdot \prod_{\tau \in \text{Pfad}(\sigma)} \frac{\alpha}{\alpha + 1} \\ &= \left(\frac{\alpha}{\alpha + 1} \right)^{|\sigma|} \\ &< \text{brT}^{-|\sigma|} \end{aligned} \quad (6.28)$$

Berücksichtigt man nun die Definition der Branching-Number aus Definition 1, so folgt

$$\inf_{\Pi} \sum_{\sigma \in \Pi} \mathbf{P}^+[\sigma \leftrightarrow r] \leq \inf_{\Pi} \sum_{\sigma \in \Pi} \text{brT}^{-|\sigma|} = 0$$

Mit Lemma 7 folgt nun für $\alpha < \alpha_{c,Perk.}$:

$$\mathbf{P}^+[\|\hat{\mathbf{T}}_r\| = \infty] = 0 \quad (6.29)$$

Wenden wir uns nun dem Fall $\alpha > \alpha_{c,Perk.}$ zu. Hier ist:

$$\frac{\alpha}{\alpha + 1} > \frac{\alpha_{c,Perk.}}{\alpha_{c,Perk.} + 1} = \max \left\{ \frac{1}{2}, \frac{1}{\text{brT}} \right\} \geq \frac{1}{\text{brT}} \quad (6.30)$$

Da wir uns in der Plus-Phase eines attraktiven Modells befinden, gilt für alle $\tau \in K_T : \mathbf{P}^+[u_\tau = +1] \geq \frac{1}{2}$, also auch $f(\tau) \geq 1$. Damit erhalten wir folgende Abschätzung:

$$\begin{aligned} & f(\tau) \cdot \alpha \geq \alpha \\ \Rightarrow & \frac{1}{f(\tau) \cdot \alpha} \leq \frac{1}{\alpha} \\ \Rightarrow & 1 + \frac{1}{f(\tau) \cdot \alpha} \leq 1 + \frac{1}{\alpha} = \frac{\alpha + 1}{\alpha} \\ \Rightarrow & \left(1 + \frac{1}{f(\tau) \cdot \alpha} \right)^{-1} \geq \frac{\alpha}{\alpha + 1} > \frac{1}{\text{brT}} \end{aligned}$$

Verwenden wir dieses Ergebnis in Gleichung (6.25), erhalten wir:

$$\begin{aligned} \mathbf{P}^+[\sigma \leftrightarrow r] &= \mathbf{P}^+[u_r = +1] \prod_{\tau \in \text{Pfad}(\sigma)} \frac{f(\tau) \cdot \alpha}{f(\tau) \cdot \alpha + 1} \\ &= \mathbf{P}^+[u_r = +1] \prod_{\tau \in \text{Pfad}(\sigma)} \left[1 + \frac{1}{f(\tau) \cdot \alpha} \right]^{-1} \\ &> \frac{1}{2} \cdot \text{brT}^{-|\sigma|} \end{aligned} \quad (6.31)$$

Unter Berücksichtigung der Definition der Branching-Number folgt nun:

$$\inf_{\Pi} \sum_{\sigma \in \Pi} \mathbf{P}^+[\sigma \leftrightarrow r] > \frac{1}{2} \inf_{\Pi} \sum_{\sigma \in \Pi} \text{br} T^{-|\sigma|} = 0$$

Daher ist es möglich, eine Zahl d so zu wählen, daß $\sum_{n \geq 1} \frac{1}{n^d} < \infty$ ist und:

$$\inf_{\Pi} \sum_{\sigma \in \Pi} \frac{1}{n^{|\sigma|}} \mathbf{P}^+[\sigma \leftrightarrow r] > 0$$

Nach Lemma 7 folgt nun:

$$\mathbf{P}^+[\|\hat{T}_r\| = \infty] > 0$$

Die fast sichere Existenz einer unendlichgroßen Zusammenhangskomponente folgt nun, da das Ereignis “ $\|\hat{T}_r\| = \infty$ ” im Tail-Field der u_σ liegt und dieses bezüglich der Plus-Phase nur Ereignisse mit Wahrscheinlichkeit 0 oder 1 beinhaltet. (vgl [7]). \square

Literaturverzeichnis

- [1] Thomas M. Cover, Joy A. Thomas, *Elements in Information Theory*, Wiley Series in Telecommunications, New York, 1991
- [2] Peter G. Doyle, J. Laurie Snell, *Random Walks and Electric Networks*, Carus Mathematical Monographs 22, Math. Assoc. Amer., Washington D.C.
- [3] Hans-Otto Georgii, *Gibbs measures and phase transition*, de Gruyter, Berlin, 1988
- [4] Harro Heuser, *Lehrbuch der Analysis, Teil 1*, B.G. Teubner, 5. Auflage, Stuttgart, 1990
- [5] Ross Kindermann, J. Laurie Snell, *Markov Random Fields and their Applications*, Contemporary Mathematics 1, American Mathematical Society
- [6] Thomas M. Liggett, *Interacting Particle Systems*, Grundlehren der mathematischen Wissenschaft 276, Springer Verlag, New York, 1985
- [7] Russel Lyons, *The Ising-Modell and Percolation on Trees and Tree-like Graphs*, Communications in Mathematical Physics 125, 1989, S. 337-355
- [8] Russel Lyons, *Random walks and percolation on trees*, Annals of Probability, 1990, Vol. 18, No.3, S. 931-958
- [9] Russel Lyons, Yuval Peres, *Probability on Trees and Networks*, Preprint vom 22.09.1998, <http://php.indiana.edu/~rdlyons/>
- [10] Christopher J. Preston, *Gibbs states on countable sets*, Camebridge University Press, London, 1974

- [11] W.Evans, C.Kenyon, Y.Peres, L.J.Schulman, *Broadcasting on trees and the Ising-Model*, Preprint vom 1. November 1998, <http://www.ma.huji.ac.il/~peres/recent.html>
- [12] David Williams, *Probability with Martingales*, Camebridge University Press, Camebridge 1991